AGARD-R-700

AGARD

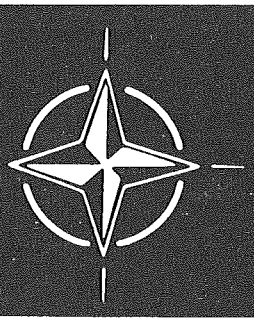
ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD REPORT No. 700

Modern Data Analysis Techniques in Noise and Vibration Problems

NORTH ATLANTIC TREATY ORGANIZATION



DISTRIBUTION AND AVAILABILITY
ON BACK COVER

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Report No.700

MODERN DATA ANALYSIS TECHNIQUES IN NOISE AND VIBRATION PROBLEMS

This Report was prepared at the request of the Fluid Dynamics and the Structures and Materials Panels of AGARD for the Special Course presented at the von Kármán Institute, Rhode-St-Genèse, Belgium on 7–11 December 1981.

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

Published November 1981

ISBN 92-835-0303-1

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London, W1P 1HD

CONTENTS

Lecture		Page
1	METHODES MODERNES D'ANALYSE ET DE TRAITEMENT DU SIGNAL UTILISEES EN ACOUSTIQUE ET VIBRATION ET, PLUS PARTICULIEREMENT, APPLIQUEES A L'AEROACOUSTIQUE par M.Perulli	1
2	FLUCTUATING STRESS FIELDS IN CONTINUOUS MEDIA by P.E.Doak	12
3, 4	VIBRATION OF STRUCTURES EXCITED ACOUSTICALLY by B.L.Clarkson	19
5	FUNDAMENTAL CONCEPTS OF SOUND RADIATION by M.C.Junger	38
6	FUNDAMENTAL CONCEPTS OF FLOW-GENERATED NOISE by W.K.Blake	43
7	EXTENDED SOUND SOURCES by M.C.Junger	54
8	STOCHASTIC EXCITATION OF ELASTIC STRUCTURES AND EXAMPLES OF FLOW-GENERATED NOISE by W.K.Blake	60
9	SOUND RADIATION BY ELASTIC STRUCTURES by M.C.Junger	72
10	THE APPLICATION OF STATISTICAL ENERGY ANALYSIS TO VIBRATION OF STRUCTURES EXCITED ACOUSTICALLY by B.L.Clarkson	76
11	LINEAR SIGNAL PROCESSING I by J.K.Hammond	80
12	LINEAR SIGNAL PROCESSING II by J.K.Hammond	92
13	NON-STATIONARITY AND NONLINEARITY IN DATA ANALYSIS by J.K.Hammond	107
14	PARAMETRIC METHODS IN SIGNAL ANALYSIS by J.K.Hammond	117
15, 16	VIBRATION TRANSMISSION AND SOUND RADIATION by M.Heckl	124
17	PROCEDURES RELATING THE NEAR- TO THE FAR-FIELD: IMAGING TECHNIQUES by M.C.Junger	142
18–1	PROGRES RECENTS DANS LA MESURE DE L'INTENSITE ACOUSTIQUE par M.Perulli	150
18–2	MODAL ANALYSIS USING DIGITAL TEST EQUIPMENT by P.Garcia	151

METHODES MODERNES D'ANALYSE ET DE TRAITEMENT DU SIGNAL UTILISEES EN ACOUSTIQUE ET VIBRATION ET, PLUS PARTICULIEREMENT, APPLIQUEES A L'AEROACOUSTIQUE

par M. PERULLI*
Chef de la Division Acoustique
Office National d'Etudes et de Recherches Aérospatiales, 92320 Châtillon - FRANCE

RESUME

Après une brève revue des procédés classiques de traitement du signal, de nouvelles méthodes sont présentées en précisant leurs bases ainsi que leur champ d'applications. Sur des exemples concrets, leur mise en oeuvre est décrite. A titre d'illustration, les interactions entre acoustique et vibrations sont exposées dans le cas du rayonnement en champ acoustique lointain.

MODERN DATA ANALYSIS TECHNIQUES IN NOISE AND VIBRATION PROBLEMS WITH PARTICULAR EMPHASIS ON AEROACOUSTIC APPLICATIONS

ABSTRACT

After a brief review of classical methods, the principles and general theorems and domains of application of modern methods of data analysis will be presented. This will be followed by details of the instrumentation requirements for the implementation of these methods and of the practical problems which arise. Finally, applications to noise and vibration problems will be considered, with reference to particular examples, many of which are chosen to illustrate the intimate connection between acoustics and vibrations in aeroacoustics.

1. INTRODUCTION

Dans les domaines de l'acoustique comme dans celui des vibrations le traitement du signal intégré dans une méthodologie de travail prend de plus en plus d'importance. En effet, l'ingénieur, pour un objectif donné, doit concevoir un schéma logique lui permettant d'atteindre cet objectif (fig. 1).

L'objectif étant fixé, le chercheur définit alors et met en oeuvre le ou les dispositifs de prises d'informations : moyens les mieux appropriés pour sonder le système à analyser. Ces moyens peuvent être :

- des transducteurs sensibles à la pression, à l'accélération, aux déplacements, à la température, etc...

- des systèmes optiques de détection : radiométrie infrarouge, mesures de vitesse par effet Doppler, strioscopie, mesures de concentration par fluorescence, mesures de densité par diffusion Rayleigh, etc...

Puis le chercheur analyse ces informations ainsi que les traitements les plus performants, compte tenu :

i) de la pauvreté des prises d'informations, c'est-à-dire de leur caractère généralement partiel, car limité dans leurs performances et lacunaires dans leur répartition spatiale ;

i^2) de la définition de niveaux de contrôle de ces informations au cours de leur traitement afin d'en assurer la validité dans le cadre de la méthodologie choisie ;

i^3) de ce que l'enchaînement des opérations permettra d'atteindre l'objectif choisi ;

i^4) des éventuels défauts de la méthodologie eu égard aux objectifs initialement fixés.

*et Professeur à l'Université de Technologie de Compiègne

Les différents exposés qui seront présentés au cours de cette semaine de travail n'aborderont pas l'aspect descriptif des moyens de prises d'informations. Néanmoins ceux-ci seront parfois présentés sommairement afin de pouvoir discuter les limitations des procédés décrits.

Par contre, l'accent sera mis sur le choix de méthodes d'analyse et de traitement du signal face à des objectifs.

Cela amènera à rappeler des notions d'acoustique et de vibrations au travers d'un souci commun : le rayonnement acoustique en champ lointain.

En effet que ce soit pour les moyens de transport dont l'aéronautique civile, les équipements militaires, l'industrie lourde, les biens de consommation, etc... ce souci est maintenant un paramètre dimensionnant pris en compte au niveau des avant-projets. Cette prise en compte est justifiée soit par la législation, en vigueur ou à venir, soit par des raisons de sécurité.

Le thème dominant de ce cours (commun à l'acoustique et aux vibrations) conduit de façon naturelle à présenter, parallèlement aux méthodologies, les aspects fondamentaux des méthodes de traitement du signal, en s'attachant à bien faire ressortir :

- leurs hypothèses de base,
- leurs champs d'application,
- leurs utilisations effectives en précisant, en particulier, les impasses que l'on est obligé de faire dans ces utilisations.

De ce fait, les séances de travail de cette semaine constituent une balance dont on trouve dans les plateaux :

- d'une part, le point de vue du chercheur attaché à des objectifs et pour lequel le traitement du signal est un outil, ce chercheur souhaitant connaître a priori la méthode de traitement ou le mode de représentation le plus performant

et

- d'autre part, le spécialiste de méthodes de traitement du signal plus rigoureux dans les principes et qui est prêt à les ajuster compte tenu des informations accessibles et des objectifs à atteindre.

Aussi ne doit-on pas s'attendre à des développements théoriques très fondamentaux et très rigoureux mais à des présentations basées sur des compromis dans lesquels le pragmatisme n'est pas exclu mais dans lesquels les fondements rigoureux restent sous-jacents.

2. MODES DE REPRESENTATION

Il est très ambitieux de vouloir effectuer une présentation exhaustive des différentes transformations auxquelles on peut soumettre un signal dans le cadre de la théorie des représentations.

A titre d'introduction à cette série d'exposés, nous allons tenter de dégager les grandes lignes des théories qui seront présentées en essayant de les classer. Pour ce faire, nous nous aidons fortement du cours de M. B. PICINBONO, [1] Professeur à l'Université d'Orsay, à l'Ecole Supérieure d'Electricité et Directeur du Laboratoire des Signaux et Systèmes.

3. CLASSIFICATION DES SIGNAUX ET DES MODES DE FILTRAGE

Les signaux que nous rencontrons généralement sont :

a) - à temps discrets ou continu. Pour les spécialistes de l'acoustique ou des vibrations, il s'agirait d'une suite de nombres semblable à celle que l'on obtient par échantillonnage de signaux continus dans le temps ;

b) - à valeurs discrètes ou continues. Il est évident que des signaux numériques sont toujours à valeurs discrètes mais on ne peut affirmer a priori, sauf redondance et précautions particulières, que des signaux à valeurs discrètes représentent des signaux à valeurs continues ;

c) - signaux déterministes à temps et valeurs continus. On sait qu'ils sont bien décrits par leur transformée de Fourier (T.F.) ou de Laplace (T.L.) c'est-à-dire qu'ils sont parfaitement compatibles avec le filtrage linéaire.

En fait, c'est le filtrage linéaire qui introduit les représentations de Fourier et de Laplace (cf. B. PICINBONO) dont les propriétés essentielles sont :

i) un filtre linéaire est un convoluteur continu ou discret,

i^2) un filtre linéaire est un système linéaire et invariant dans le temps,

i^3) les filtres linéaires sont des systèmes linéaires dont les fonctions propres sont les signaux exponentiels : ce résultat est d'une importance pratique capitale.

De ces propriétés on déduit :

- la fonction de transfert du filtre H(p) qui est reliée à la transformée de Laplace bilatérale R(t),

- le gain complexe du filtre linéaire $H(2 i \Pi \nu) = G(\nu)$,

- les fonctions propres des filtres linéaires qui constituent une base pour le développement des signaux.

Il est à remarquer que dans le cas de signaux discrets, le filtrage linéaire discret est un convoluteur discret et les fonctions propres de ce filtrage sont les exponentielles z^n ; la transformée z (T.z) est alors définie par la suite discrète :

$$\hat{x}(z) = \sum_n x_n z^{-n}$$

Dans le cas de ces signaux déterministes à temps et valeurs continus, une propriété fondamentale pour les applications est à rappeler : la T.F. d'une fonction réelle x(t) possède une symétrie hermitienne ce qui entraîne que les fréquences négatives n'apportent aucune information nouvelle sur $x(\nu)$. On définit le filtre en quadrature F_Q qui associe à tout cosinus le sinus de même fréquence et on passe de x(t) à y(t) par la transformation de Hilbert définie par :

$y(t) = f \frac{x(\theta)}{\Pi(t - \theta)} d\theta$, sachant que $x(t) = f \frac{y(\theta)}{\Pi(t - \theta)} d\theta$.

On peut alors définir un filtre analytique F_A par : $F_A = F_I + i F_Q$ qui permet de construire le signal analytique (S.A.) z(t) associé au signal réel x(t). On montre alors que les parties réelles et imaginaires du S.A. sont transformées de Hilbert l'une de l'autre. Cette propriété se retrouve dans l'étude de la causalité : ce qui définit un critère pratique caractérisant un signal donc, par extension, le milieu émettant ce signal. (Rappelons que si le signal x(t) est nul pour $t < 0$, on dit qu'il est causal. Dans ce cas sa T.L. est monolatérale et la T.L. bilatérale est équivalente à la T.F.).

Mentionnons, de plus, que pour ces signaux on peut également utiliser des filtres dynamiques, des filtres numériques et des filtres récursifs. Les théories correspondantes seront présentées par ailleurs.

d) - Signaux aléatoires à temps continus et discrets. Nous savons que le signal aléatoire se différencie du signal déterministe par le fait qu'il a une forme imprévisible à l'avance et qu'il ne peut être décrit par une fonction mathématique x(t) connue.

De ce fait et par usage (cf. le cours de M. B. PICINBONO), on introduit des notions de calcul des probabilités et de variable aléatoire (V.A.). J.K. HAMMOND présentera de façon plus rigoureuse les notions de base associées à ces signaux, sachant que l'on transpose au cas des signaux aléatoires les notions d'analyse utilisées pour des signaux déterministes. De plus, on considère des propriétés spécifiques sans lesquelles le traitement de ces signaux se heurterait à de sérieuses difficultés :

i) la stationnarité : cette propriété des signaux aléatoires (S.A.) implique que certaines de leurs propriétés statistiques sont invariantes dans toutes translations de l'origine du temps. On peut montrer que, sous certaines hypothèses, cette propriété se conserve par filtrage linéaire ;

i2) l'ergodisme : cette propriété des S.A. est beaucoup plus subtile. Elle équivaut, en gros, à considérer que des moyennes d'ensemble, en général inaccessibles, sont, dans la pratique, approximables par des moyennes temporelles prises sur une épreuve, tout au long de l'histoire du mécanisme considéré.

Suite à ces quelques rappels de base qui conduisent aux notions classiques de corrélations, de moments, de densités spectrales pour lesquelles la notion de filtrage linéaire reste valable, nous allons, à l'aide d'exemples concrets, montrer d'autres développements dont les prémices sont prometteuses.

4. SIGNATURES COMPAREES

Une première idée consiste à se demander dans quelle mesure la signature acoustique d'une distribution circulaire de sources monopolaires cohérentes en phase, diffère de celle d'une distribution incohérente.

Cette situation se trouve pratiquement dans le cas d'écoulements libres issus d'une tuyère. En effet, la zone de mélange se compose d'une zone laminaire (cône potentiel) entourée d'une couronne circulaire fluctuante dans laquelle se superposent, d'après l'expérience et certaines théories, des structures déterministes et des structures aléatoires. La question fondamentale qui se pose, en vue de la réduction du champ sonore lointain, est la contribution sonore relative de ces deux classes de structures qui correspondent à des mécanismes physiques différents.

Une amorce de réponse est apportée à l'aide de mesures circulaires effectuées en champ lointain par deux microphones. En effet, l'un étant fixe, l'autre mobile sur un cercle centré sur l'axe de l'écoulement et situé dans le plan des sources, on calcule la fonction d'intercorrélation spatiale qui ne dépend que de l'angle ψ séparant les deux

microphones [2]. On constate la quasi-impossibilité de distinguer la signature de ces deux sortes de structure dès que le nombre des modes azimutaux, existant simultanément dans les sources cohérentes, est supérieur à 6.

La figure 2 est relative à un nombre de Strouhal $fD/C = (1/\Pi)\ \omega\ r_0/C \simeq 0{,}32$ et la figure 3, au nombre de Strouhal Str ≃ 1,6. Il est à remarquer que cette propriété est conservée en champ proche : figure 4 pour Str ≃ 0,32 et figure 5 pour Str ≃ 1,6.

Ces résultats montrent les difficultés d'une interprétation physique des résultats de mesure dont les signaux ont été traités suivant une certaine logique : système cylindrique d'où corrélations polaires. Seulement cette logique conduit, dans ce cas, à une impasse.

5. SIGNATURE ACOUSTIQUE D'UN PAQUET D'ONDES

Prenons comme objectif d'une étude de bruit d'écoulements libres, la confrontation aux modélisations théoriques (cf. fig. 1).

5.1 - La modélisation d'écoulements libres subsoniques repose essentiellement sur des concepts de turbulence stationnaire. De ce fait, des corrélations spatio-temporelles effectuées dans le volume source de bruit devraient, par double affinité, conduire à une courbe unique étroitement liée à la fonction de corrélation de ce volume source.

La figure 6, qui n'est qu'un exemple parmi d'autres, montre qu'il existe, d'une part, une dispersion non négligeable des temps et des amplitudes d'où une non-stationnarité du milieu et, d'autre part, la présence de parties négatives dont l'amplitude non négligeable laisse présager que l'on se trouve en présence d'un problème multi-échelles.

5.2 - Dans le but d'effectuer une bonne séparation d'échelles, nous avons développé des techniques de traitement du signal faisant appel au filtrage non linéaire et basées sur l'échantillonnage conditionnel [3]. Elles nous permettent de suivre l'évolution de paquets d'ondes définis par des critères d'amplitude et de phase, de telle sorte que l'on puisse en prévoir ou en montrer les conséquences acoustiques [4, 5, 6], avec ou sans recouvrement fréquentiel partiel dû à des mécanismes différents.

Pour illustrer cette séparation d'échelles, nous présentons des vues obtenues dans un jet libre chaud de forte vitesse à l'aide de la strioscopie conditionnelle à couteau (fig. 7). Cette méthode équivaut à effectuer de la stroboscopie à partir des conditions (amplitude et phase) appliquées à un signal (issu d'un radiomètre, d'un microphone, d'un fil chaud ou d'un anémomètre laser). Ainsi en superposant les images obtenues avec quelques centaines d'éclairs, on isole le phénomène choisi (c'est-à-dire que l'on améliore le rapport signal sur bruit au sens du contraste) (fig. 8), puis à l'aide d'un dispositif électronique on peut faire se déplacer le phénomène visualisé et définir ainsi son domaine d'existence. Sur la figure 9, trois images sont présentées :

a) la structure est assez nette ;

b) on l'a déplacée arbitrairement et elle commence à se dégrader ;

c) elle se dégrade de plus en plus pour disparaître.

Connaissant les conditions de traitement du signal, on peut déduire la vitesse à laquelle se déplace le phénomène ainsi isolé et son extension spatiale.

5.3 - Dans l'exemple présenté ci-dessus, l'information microphonique quantifiée (en amplitude et phase) permet d'y associer un mécanisme physique dans le volume source de bruit. Lorsque celui-ci n'est pas accessible on peut effectuer un relevé d'informations de pression instationnaire en champ proche sous forme d'une matrice plane dont on peut déduire des lignes d'isophase et d'iso-amplitude.

Avec le souci précédent de connaître l'efficacité acoustique d'un paquet d'ondes on peut, de plus, déduire ces lignes isophases et iso-amplitudes corrélativement à un petit élément de volume de l'écoulement libre. Un exemple est présenté sur les figures 10 et 11 :

- la figure 10 représente des lignes iso-amplitude dont la directivité semble indiquer que l'émission maximale serait proche du plan de sortie de la tuyère ;

- la figure 11 montre des lignes isophases. Il est à mentionner que la distance entre deux isophases dont la différence est égale à 2Π est compatible avec la célérité du son à l'infini.

On déduit ce remarquable résultat : en terme de longueurs d'ondes rapportées au diamètre de la tuyère, le champ proche contient des paquets d'ondes dont une partie de son énergie est véhiculée à la vitesse du son, donc est propagative.

6. TRAITEMENT SYNCHRONE NON LINEAIRE

Dans le cas de systèmes tournants en mouvement, le traitement synchrone non linéaire améliore de façon remarquable le rapport signal sur bruit et surtout facilite la confrontation avec la théorie.

6.1 - Pour des systèmes tournants au point fixe, le dépouillement des mesures en synchronisme avec la rotation est maintenant classique et fournit des informations, en général, simples d'interprétation et inaccessibles autrement.

Dans le cas de la propagation du son dans des conduits, la mesure des spectres de nombres d'ondes a quitté le domaine du laboratoire pour être aisément utilisée par les industriels. Il est ainsi possible de vérifier la relation de dispersion et/ou des codes de calcul de propagation en milieux hétérogènes (cf.[7]par exemple).

6.2 - Pour des systèmes tournants en mouvement il n'en est plus de même. Prenons le cas d'une mesure de bruit de survol d'avion de tourisme [8].

Entre la source de bruit et le point de mesure au sol, les ondes sonores sont dispersées et absorbées par l'atmosphère, puis se réfléchissent au sol. Cette dernière phase, à laquelle nous nous intéresserons ici, peut modifier fortement le bruit perçu par un microphone situé à 1,2 m du sol, par rapport à ce qu'il serait en espace libre.

6.2.1 - Réflexions au sol

L'avion effectuant son survol à une altitude de 1000 pieds, les ondes acoustiques qui atteignent le sol peuvent être représentées par des ondes planes. Au voisinage du sol, il se crée un système d'interférences et la pression à une hauteur h au-dessus du sol résulte de la superposition du champ direct et du champ réfléchi.

Il existe de nombreux modèles pour représenter les propriétés de réflexion du sol, mais nous n'avons retenu, dans cette première approche, que le modèle le plus simple : le sol a une impédance réelle.

Les interférences entre le champ direct et le champ réfléchi se caractérisent localement par l'existence de fréquences pour lesquelles l'interférence est totale (fréquences d'extinction).

Il est classique que les fréquences d'extinction (fig. 12) pour une onde plane incidente sous un angle α, soient prévues par la formule $f_d = \dfrac{(2n + 1)c}{4h \sin\alpha}$ c'est-à-dire qu'elles se présentent comme des "trous" dans le spectre de bruit reçu, à des fréquences qui sont des multiples d'une fréquence dite fréquence fondamentale d'extinction du signal (fig. 13). Cette dernière est fonction de la position de l'avion à l'instant où les ondes ont été émises et évolue au cours du survol.

Cette évolution, calculée théoriquement, peut se comparer à ce que l'on déduit des mesures de survol de la façon suivante : prenant un enregistrement de bruit, qui dure environ 30 s, on fait une analyse spectrale du signal toutes les secondes, et on relève sur chaque spectre la fréquence d'extinction fondamentale que l'on compare aux valeurs prédites. Malgré le faible temps d'intégration (1 s), l'évolution rapide des spectres en niveaux, due à la directivité de la source, et le déplacement des raies, dû à l'effet Doppler, on arrive assez bien à suivre le mouvement de l'appareil.

De la comparaison des maxima et des minima successifs des niveaux sur les spectres, on peut également déduire, dans le cadre de ce modèle, la valeur du coefficient de réflexion r.

Celui-ci est donné [9], en fonction des niveaux de pression p(f) par :

$$\frac{(1 + r)^2}{(1 - r)^2} = \frac{|p(2nf_d)|^2}{|p(2n + 1)f_d|} \quad .$$

6.2.2 - Analyse de survols : traitement du signal

a) - Principe

L'idée générale est d'éliminer le glissement de fréquences dû à l'effet Doppler créé par le mouvement de l'appareil en pilotant l'analyse de Fourier du signal par une fréquence multiple de la source corrigée de l'effet Doppler. Le problème principal soulevé par cette procédure, que l'on trouve évoquée dans la littérature, est celui de la détection de la fréquence de pilotage.

Certains auteurs [10] ont proposé de procéder à une trajectographie optique, de déterminer par le calcul, connaissant la position et la vitesse de l'avion, la fréquence que l'on devrait théoriquement percevoir, et d'asservir la numérisation à cette fréquence. Il est d'un emploi beaucoup plus souple, et c'est la solution que nous avons retenue, d'asservir la numérisation au signal acoustique lui-même (trajectographie acoustique). Les difficultés soulevées par cette technique et les procédures utilisées pour les surmonter sont détaillées en [11].

b) - Résultats de l'analyse

Les traitements ont été faits sur des enregistrements de bruit de survol communiqués par le Service Technique de la Navigation Aérienne (STNA) et fournissent :

- l'évolution au cours du temps de survol, d'une fréquence caractéristique de l'appareil (fig. 14) ;

- l'évolution au cours du temps des niveaux sur les raies que l'on désire analyser (fig. 15 et 16).

Si f est la fréquence émise par la source, M le nombre de Mach d'avancement, la fréquence perçue lorsque l'avion est à l'infini amont est :

$f_1 = \frac{f}{1 - M}$ et $f_2 = \frac{f}{1 + M}$ lorsque l'avion est à l'infini aval.

On en déduit : $M = \frac{f_1 - f_2}{f_1 + f_2}$; $f = \frac{2 f_1 f_2}{f_1 + f_2}$.

Ces résultats sont en bonne concordance avec les conditions de vol affichées.

7. CONCLUSIONS

Ces quelques exemples de traitement du signal au sens :

- d'une numérisation non linéaire des informations afin de s'affranchir (ou d'inclure) de l'effet Doppler lié au mouvement du système,

- d'un filtrage non linéaire dans le but de privilégier un mécanisme physique particulier,

montre que le champ d'investigation est vaste d'autant que, dans cette présentation introductive, nous avons volontairement omis d'autres techniques de traitement :

- l'analyse spectrale,
- l'analyse factorielle,

des techniques de mesure à traitement du signal intégré :

- l'intensimétrie sonore,
- l'holographie acoustique dans le spectre audible.

Ces techniques feront l'objet de présentations spécifiques.

De même, des exposés seront consacrés à la définition de critères de rayonnement en champ lointain qui complètent celui présenté au paragraphe 5.3 ci-dessus.

8. ELEMENTS DE BIBLIOGRAPHIE

[1] - B. PICINBONO, Eléments de théorie du signal, Dunod Université, (1978).

[2] - Ch. BONNET, Correlation techniques and modal decomposition analysis for the detection of azimuthally coherent structures in jet flow, M. Sc de l'Université de Southampton (1979).

[3] - J. LAUFER and M.A. BADRI NARAYANAN, Mean period of the turbulent production mechanism in a boundary layer. Phys. Fluids, 14, 182 (1971).

[4] - C. DAHAN, G. ELIAS, J. MAULARD and M. PERULLI, Coherent structures in the mixing zone of a subsonic hot free jet, dans "Structure and Mechanisms of Turbulence II", Lecture Notes in Physics, n° 76 (1978).

[5] - L. AVEZARD, C. DAHAN, G. ELIAS, A. LELARGE, and J. MAULARD, Simultaneous characterization of jet noise sources and acoustic field by a new application of conditional sampling. AIAA Paper 77-1349, AIAA Journal, 16, 1121 (1978).

[6] - C. DAHAN, G. ELIAS, J. MAULARD and M. PERULLI, Coherent structures in the mixing zone of a subsonic hot free jet. J.S.V. 59, 313 (1978).

[7] - J. TAILLET, M. PERULLI, S. LEWY et J.L. PRIEUR, Répartition de pression instationnaire et propagation du bruit dans les manches d'entrée des turbomachines. L'Aéronautique et l'Astronautique, n° 63, 1977-2, p. 43.

[8] C. DAHAN, L. AVEZARD, G. GUILLIEN, C. MALARMEY et J. CHOMBARD, Bruit des avions légers à hélice. La Recherche Aérospatiale 1979-2, p. 141-149, TP n° 1979-56.

[9] - E.D. GRIFFITH and J.D. REVELL, Low noise propeller technology. AFAPL TR 73.115 (1973).

[10] - F.W.J. VAN DEVENTER, Review of aircraft noise research at Delf University of Technology. Gart-Eur 5 (1977).

[11] - C. DAHAN, A. JULIENNE et C. MALARMEY, Traitement des signaux faiblement instationnaires. Document ONERA non publié.

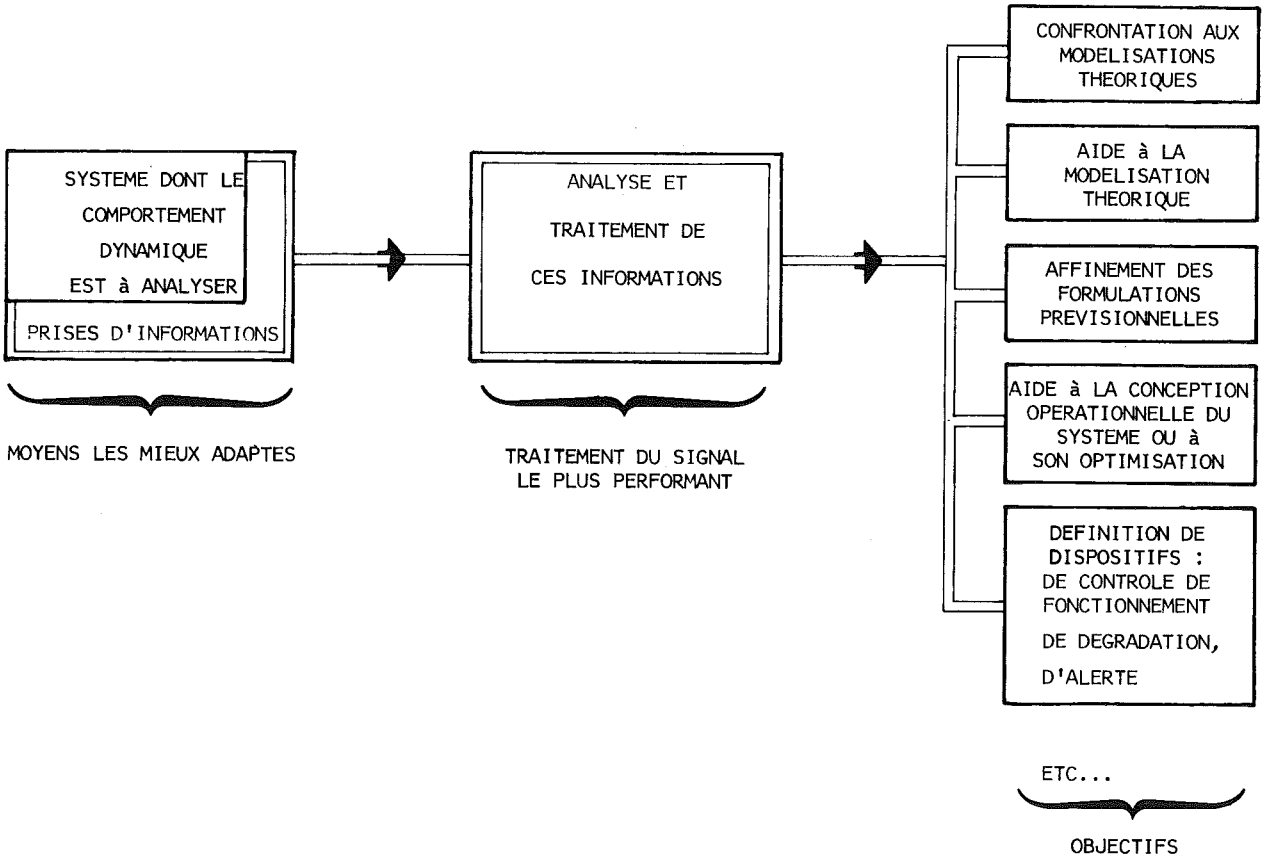


Fig. 1 — Méthodologie classique.

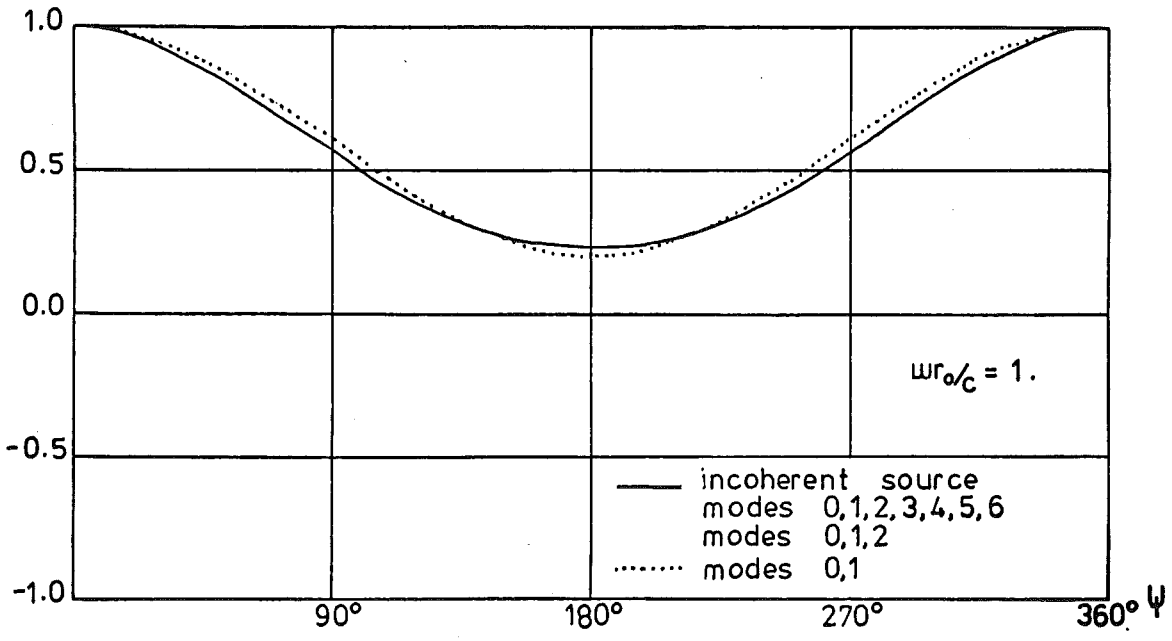


Fig. 2 — Comparaison des signatures polaires de sources monopolaires circulaires spatialement cohérentes et incohérentes au nombre de Strouhal : $Str = fD/C = (1/\pi)\ \omega r_0/C = 0{,}32$.

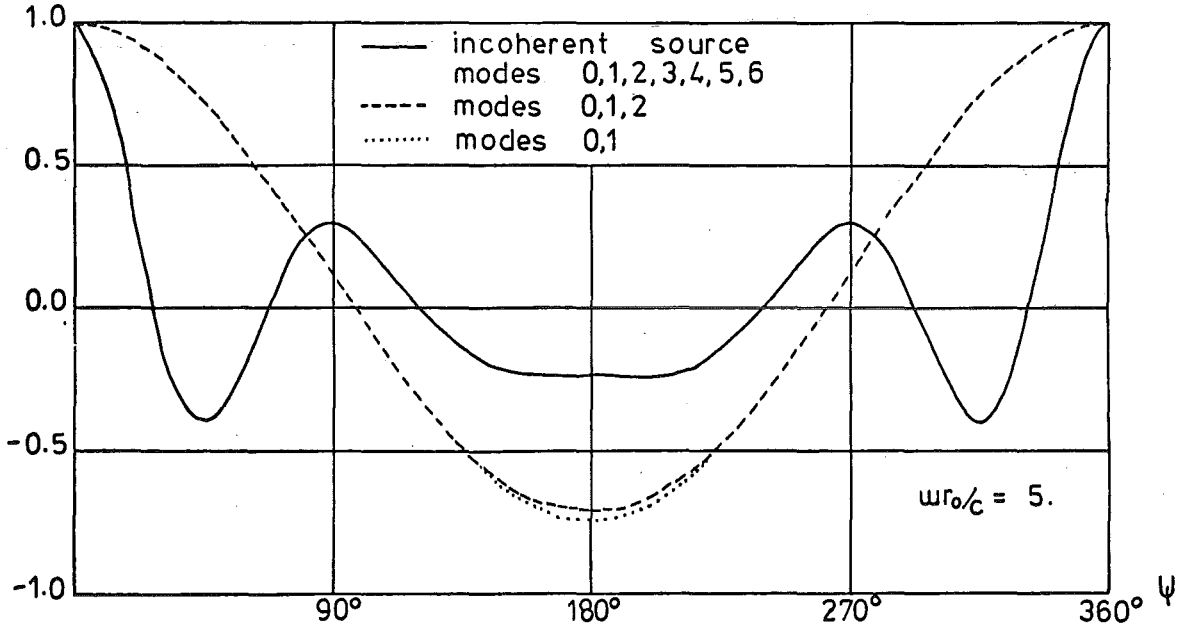


Fig. 3 — Comparaison des signatures polaires de sources monopolaires circulaires spatialement cohérentes et incohérentes au nombre de Strouhal : $Str = fD/C = (1/\pi)\ \omega r_o/c \simeq 1{,}6$.

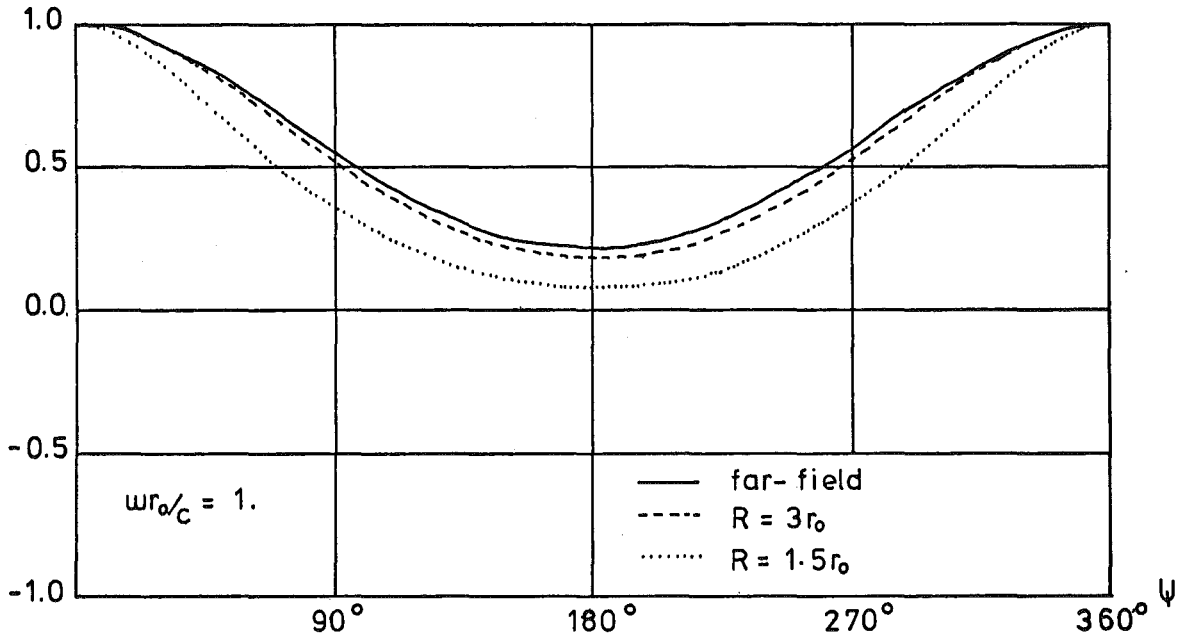


Fig. 4 — Fonctions d'intercorrélation polaire en champ proche pour $Str \simeq 0{,}32$.

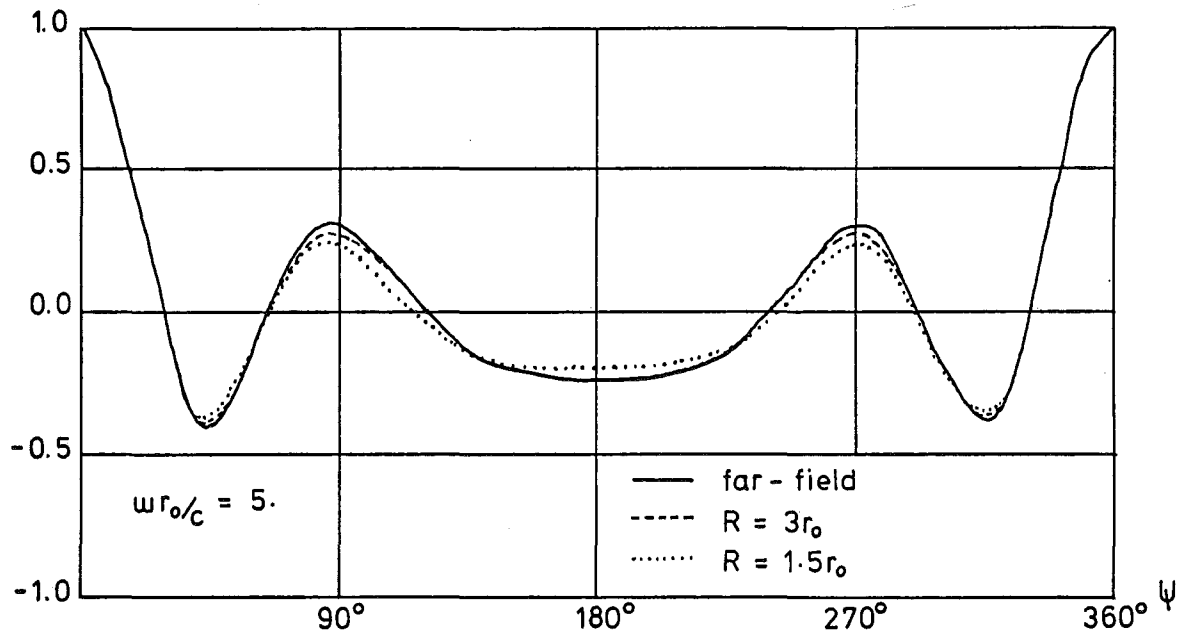


Fig. 5 — Fonctions d'intercorrélation polaire en champ proche pour Str ≃ 1,6.

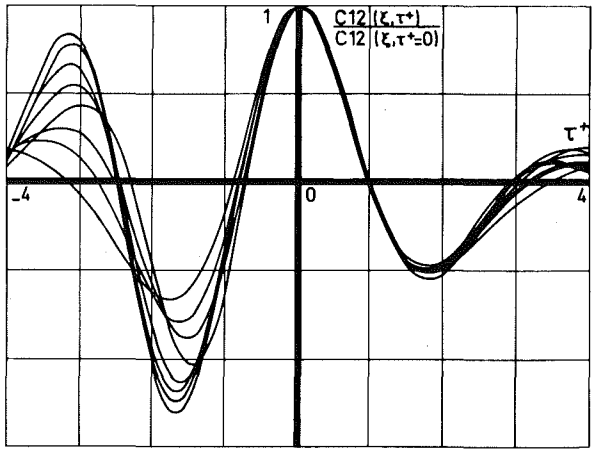


Fig. 6 — Superposition par double affinité d'intercorrélations spatio-temporelle. D = 80 mm ; Tj = 900 K ; Vj = 410 ms⁻¹ ; X/D₁ = 1 ; r/D = 0,5 ; ξ/D de 0 à 3 ; $\tau^+ = (\tau - \tau_{max})/(\tau_0 - \tau_{max})$

— τ_{max} = retard incremental pour $C_{12\,max}$.
— τ_0 = retard incremental pour le premier zéro de C_{12}.

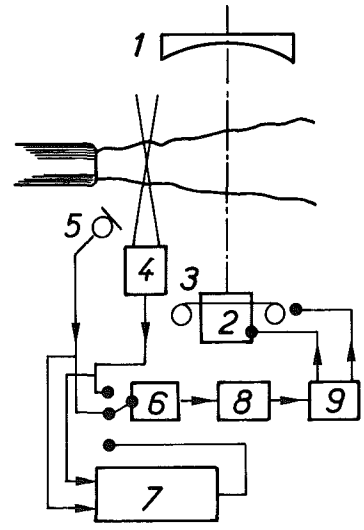


Fig. 7 — Dispositif de strioscopie conditionnelle.

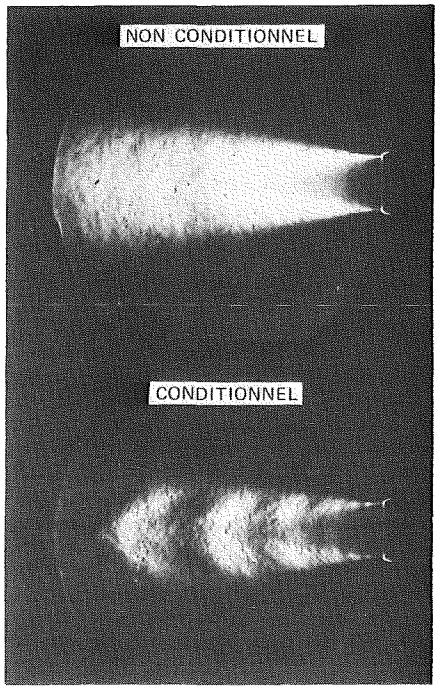


Fig. 8 — Visualisation par strioscopie conditionnelle.

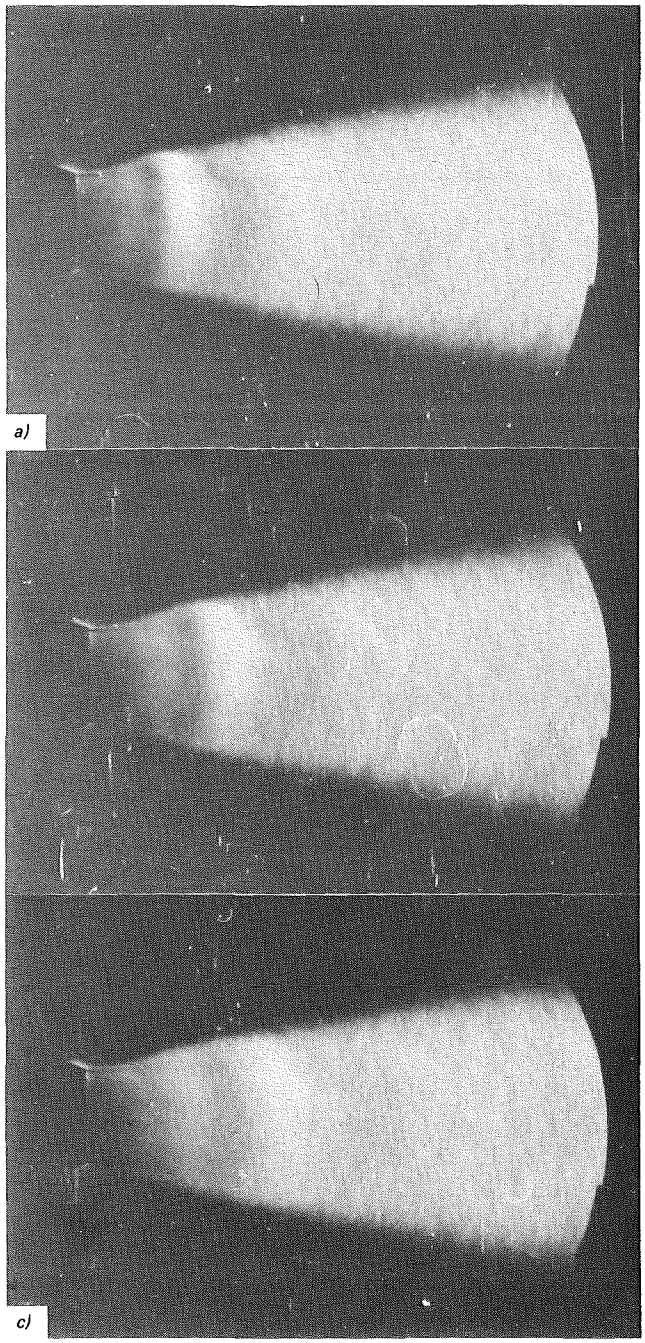


Fig. 9 — Evolution des structures visualisées.

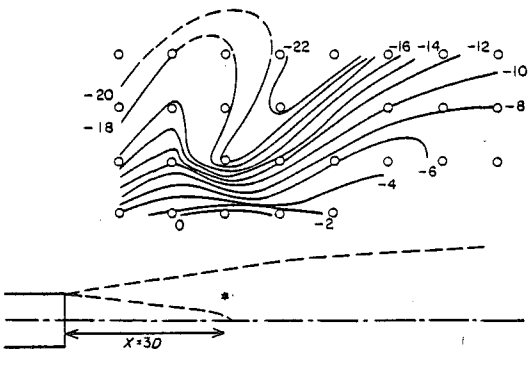


Fig. 10 — Lignes d'iso-amplitude entre un radiomètre et les microphones.

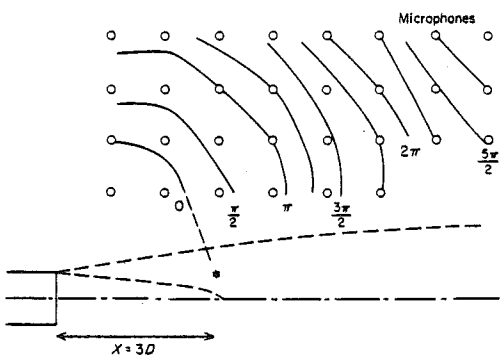


Fig. 11 — Lignes isophases correspondant à la figure 10.

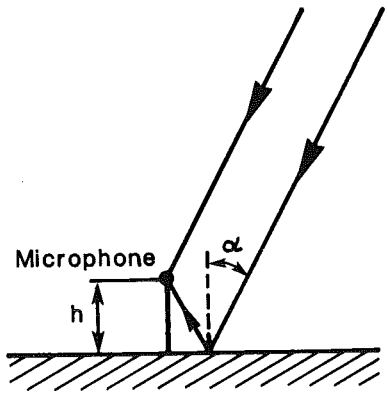


Fig. 12 – Géométrie des réflexions au sol.

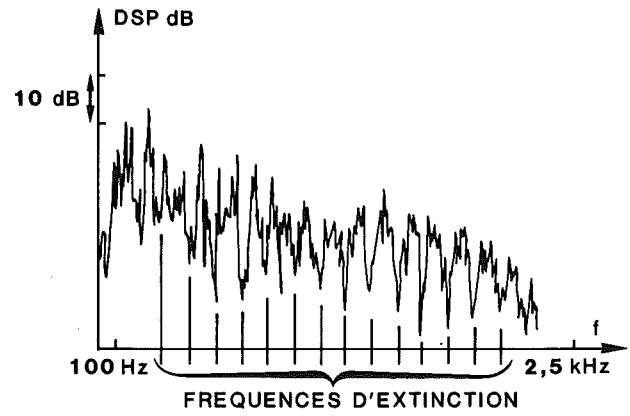


Fig. 13 – Spectre instantané de bruit de survol.

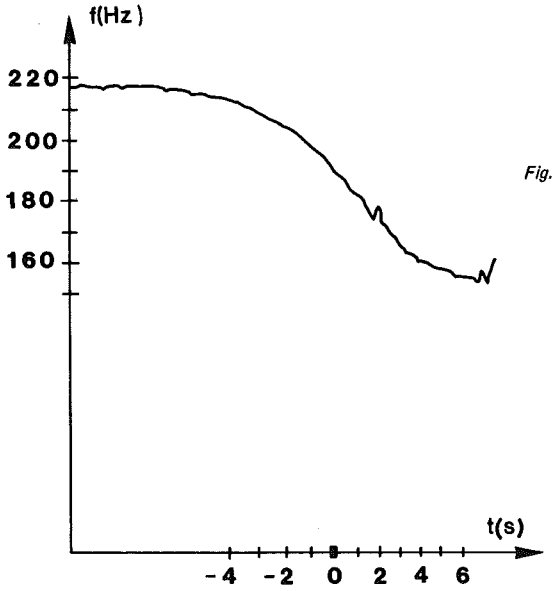


Fig. 14 – Evolution au cours du temps d'une fréquence caractéristique de l'appareil.

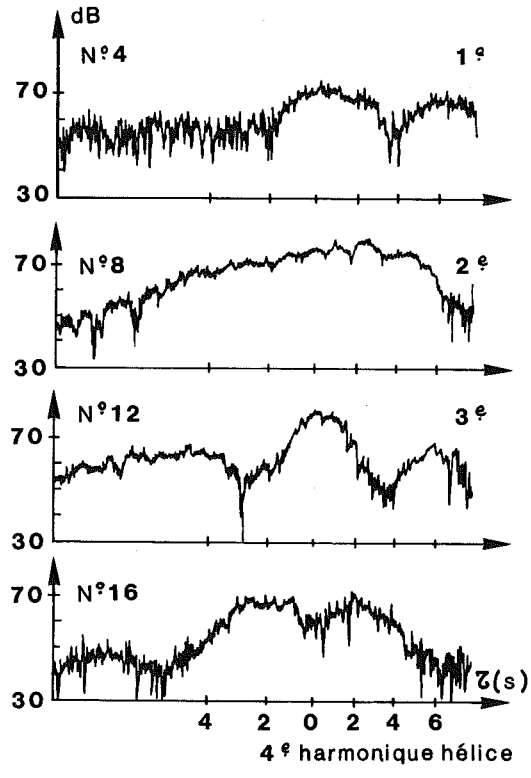


Fig. 15 – Evolution de niveaux d'harmoniques du bruit d'hélice.

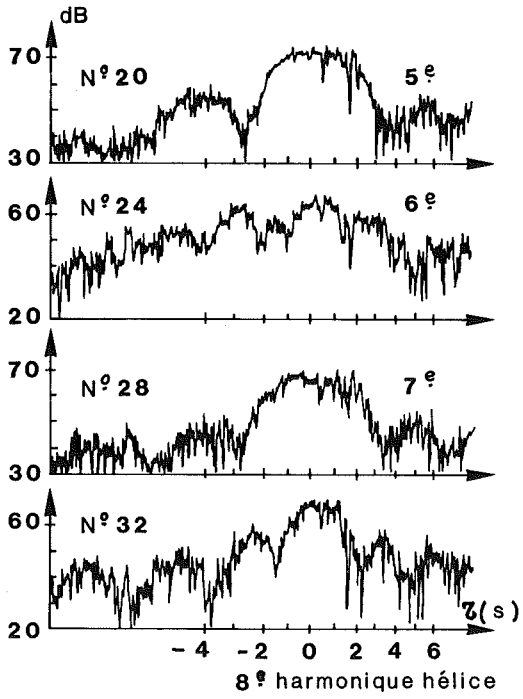


Fig. 16 – Evolution de niveaux d'harmoniques du bruit d'hélice.

FLUCTUATING STRESS FIELDS IN CONTINUOUS MEDIA

P.E. Doak
British Aerospace Professor of Acoustics
Institute of Sound and Vibration Research
University of Southampton
Southampton SO9 5NH
U.K.

In noise and vibration problems, the basic physical phenomena are those of stress wave fields in materials. The fundamental nature and behaviour of these fields in real materials, and in particular any aspects thereof that are common to materials in general, thus are of especial importance in connection with questions of data analysis techniques, both because the development and use of such techniques is in response to demands for ways of measuring these basic physical phenomena, and not just done for its own sake, and because in real material systems these basic stress wave phenomena are usually very complicated, so that the type of technique has to be carefully selected and interpreted in order that useful rather than confusing information is obtained.

The most important aspect of fluctuating stress wave fields in engineering systems of materials, in this connection, is that more often than not the mechanical behaviour is linear, or nearly so: in other words, the principle of superposition is usually applicable. Even when physical non-linearity is significant, the notorious difficulty and complexity of full non-linear theory is such that, for engineering purposes, it is usually preferable to think in terms of theoretical models that are "piecewise linear", or at least to use "mathematically linear" tools, such as Galerkin-type methods involving linear superposition of orthogonal "shape functions", etc. The well known Lighthill "acoustic analogy" used in aeroacoustics theory in fact is an important case in which all the non-linearity is consigned into a supposedly known "equivalent source" function driving an "equivalent" linear system; in this case "piecewise linearity" has been taken to the limit, with all the non-linearity being taken out of the analysis problem itself, and with quite successful results.

The finite difference forms of the partial differential equations for linear continuous mechanical systems correspond to those for three-dimensional lumped parameter networks, the lumped elements including masses, springs, and dashpots (or, in the electrical network analogy, inductances, inverse capacitances, and resistances). Such lumped parameter networks are often not directly useful in analysis, because the systems of interest are seldom compact enough to allow representation with a reasonably small number of degrees of freedom. This correspondence is conceptually useful, however, and often of indirect analytical use, because of the clear one-to-one correspondence between the denumerable eigenvalues and eigenfunctions of the continuum representation (the first N of those, say) and the N eigenvalues and eigenvectors, respectively, of the finite difference representation. Indeed, as reference to Volume I, Chapter VI of Lord Rayleigh's Theory of Sound shows, consideration of the correspondence can lead directly to formulation of the Rayleigh-Ritz method (and hence Galerkin, etc., variational and weighted residual methods) for approximate solution of eigenvalue problems. Also, the book by L. Brillouin, Wave Propagation in Periodic Structures, is, almost in its entirety, an example of the fruitfulness of this correspondence (and thinking in terms of networks) as a basis for developing useful analytical tools.

Another aspect of this correspondence of some interest is that it is relatively easy to understand how complicated networks, in particular cases, can be approximated by simpler ones, and in this way some insight can be gained into how approximate partial differential equations can be devised to replace the full elasticity equations representations of structures such as beams, plates and shells, which are "acoustically compact" (i.e., small compared to relevant wavelengths of the continuum elastic stress waves) in one or more dimensions.

The homogeneous isotropic Hookean (stress proportional to strain) continuum is perhaps the simplest ideal continuum model of considerable generality in both theory and application. It can provide a good representation of small amplitude stress wave fields in solids, and in fluids when the shear stress coefficients are taken to be negligible. The Kelvin-Voigt thermoviscoelastic continuum model allows one to include viscous and thermal stresses as well, with addition of stresses proportional to rates of strain and temperature, as well as to strain. This continuum model thus includes, as special cases, both Stokesian (viscous and thermally conducting) fluids and Hookean solids, and, independently, Cauchy thermoelastic materials. A reliable, self-consistent basic description of these, and other useful continuum models, is contained in the book by A.C. Eringen, Mechanics of Continua.

Although to date it has been applied almost exclusively to aeroacoustics, Lighthill's "acoustic analogy" is applicable, in a formal way, to any continuum, and this has consequences relevant to the analysis of stress wave fields in general. For any mass conserving continuum, the equation of mass transport is

$$\partial\rho/\partial t + \partial m_i/\partial x_i = 0, \tag{1}$$

where ρ is the mass density and m_i is the linear momentum density, $m_i \equiv \rho v_i$, v_i being the barycentric (local centre of mass) velocity. The equation of linear momentum transport is

$$\partial m_i/\partial t + \partial p/\partial x_i + \partial\{(m_i m_j/\rho) + \sigma_{ij}\}/\partial x_j = \rho f_i, \tag{2}$$

where p is the thermodynamic pressure, $\sigma_{ij} \equiv p_{ij} - p\delta_{ij}$ is the stress tensor p_{ij} less the isotropic thermodynamic pressure stress, δ_{ij} being the Kronecker delta, and f_i is the external force per unit mass (due to gravitational and/or electromagnetic force fields). Elimination of the linear terms in m_i between equations (1) and (2) gives

$$\partial^2 p/\partial x_i^2 - \partial^2 \rho/\partial t^2 = \partial(\rho f_i)/\partial x_i - \partial^2\{(m_i m_j/\rho) + \sigma_{ij}\}/\partial x_i \partial x_j. \quad (3)$$

On the reasonable assumption that a thermodynamic constitutive equation relating ρ and other thermodynamic variables is available, of the form, say, $\rho = \rho(p,S,\ldots)$, where $(\ldots)$ represents the possibility that there may be independent thermodynamic (i.e., "internal energy") variables in addition to p and the entropy S, and with the concise notation ρ_S meaning $(\partial\rho/\partial S)_{p,\ldots}$, the subscript p,... meaning that p and the other thermodynamic variables are constant, and in particular $\rho_p \equiv (\partial\rho/\partial p)_{S,\ldots} \equiv 1/c^2$, one has

$$\partial^2 p/\partial x_i^2 - \partial\{(1/c^2)\,\partial p/\partial t\} = \partial(\rho_{S,\ldots}\,\partial S/\partial t)/\partial t + (\ldots) + \partial(\rho f_i)/\partial x_i - \partial^2\{(m_i m_j/\rho) + \sigma_{ij}\}/\partial x_i \partial x_j, \quad (4)$$

which is evidently an inhomogeneous scalar wave equation for the normal stress p, with a variable wave speed c, and "equivalent source distribution" terms on the right-hand side.

Similarly, if one takes the curl of equation (2), one obtains (in mixed vector notation)

$$\partial(\text{curl } \vec{m})_i/\partial t + (\text{curl}\{\partial(\sigma_{ij} + m_i m_j/\rho)/\partial x_j\})_i = (\text{curl } \rho\vec{f})_i. \quad (5)$$

For a Hookean continuum in small amplitude motion about a uniform equilibrium rest state $(\text{curl}(\partial\sigma_{ij}/\partial x_j))_i$ would be simply $\rho_o c_{s,o}^2 (\text{curl curl curl } \vec{\xi})_i$ and $(\text{curl } \vec{m})_i$ would be $\rho_o (\text{curl } \partial\vec{\xi}/\partial t)_i$, where a subscript zero denotes a constant equilibrium value, $c_{s,o}$ is the equilibrium speed of transverse shear stress waves and $\vec{\xi}$ is the barycentric material displacement. Hence the solenoidal component of the displacement (and/or the shear stress) satisfies an inhomogeneous vector wave equation

$$\partial\{(1/c_{s,o}^2)\,\partial\xi_i/\partial t\}/\partial t - \partial^2\xi_i/\partial x_j^2 = \chi_i, \quad (6)$$

χ_i being the solenoidal vector $\text{curl}(\rho f_i)/\rho_o c_{s,o}^2$. A formally similar result, with a variable shear wave speed and a differently defined source term evidently can be obtained in cases when the material is not of simple Hookean type.

Hence the normal stress, and the Cartesian components of the shear stress, for <u>any</u> continuum, all can be regarded as satisfying forms of the inhomogeneous scalar wave equations, with a variable wave speed and a suitably defined "equivalent source distribution". These are useful insofar as the "equivalent source distribution" can be specified, even if only approximately, independently of knowledge of the dependent stress variable itself, which is to be determined by solution of the equation. The many applications of equation (4) to aeroacoustics, and the cited example of a Hookean solid, together with consideration of the forms of these "equivalent source distribution" terms in more general cases (for example, Lord Rayleigh pointed out how scattering problems could be modelled in this way) indicate that at least approximate specification of these terms is possible in a very wide variety of situations.

When the energy transport equation is considered, along with the mass and momentum equations (1) and (2), it turns out that in addition to the irrotational normal stress and solenoidal shear stress motions, a thermal diffusion motion may also occur, with the entropy, say, as the dependent variable, satisfying an inhomogeneous scalar diffusion equation of the form

$$\partial^2 S/\partial x_i^2 - (1/\kappa_S)\,\partial S/\partial t = -q_S(x_k,t),$$

where κ_S is the coefficient of thermal conductivity and $q_S(x_k,t)$ is an equivalent (non-linear) entropy source distribution.

When <u>gradients</u> of one or another of mean quantities (such as stresses, barycentric velocities, temperatures, etc.) are not suitably small, the source terms of the resulting scalar wave, vector wave and diffusion equations, respectively, contain terms that are linear in the other fluctuations: i.e., the source term for the pressure has terms linear in the shear stress and thermal fluctuations, etc. Thus the normal stress, shear stress and thermal stress motions are then linearly coupled in proportion to the magnitudes of these mean gradients. The source term for the pressure fluctuations p, for example, can still be specified independently of knowledge of p, to first order, but not without knowledge to first order of the shear stress and entropy fluctuations.

Finally, when mean motion is appreciable, convection of the stress and thermal fluctuations with the material motion occurs, and thus it may be desirable, in the inhomogeneous wave and diffusion equations, to transfer some right-hand side "equivalent source" terms to the left-hand side to represent these mean convection effects as distinct from "generation" effects, these terms being linear in the respective fluctuating quantities and their derivatives but having spatially dependent coefficients, known when the mean velocity field is known.

Thus, with considerable generality, acoustic and vibration problems can be regarded as those of the solution of sets of inhomogeneous scalar wave and diffusion equations, perhaps of convected form and perhaps linearly coupled through their respective "equivalent source" terms, together with appropriate boundary conditions on the surfaces separating their respective domains of validity.

The wave speeds in these equations are in general functions of the thermodynamic variables, and hence not even the homogeneous forms of the equations are linear, unless the fluctuations in the wave speeds can be neglected and their spatial dependence can be specified (through knowledge of the mean temperature, say). Again, it is very often reasonable to make such approximations and specifications, and then one has equations of the form, written here specifically for the pressure fluctuations (as an example), with

$p(x_k,t)$ now assumed to represent fluctuations of zero time average,

$$c^2(x_k)\,\partial^2 p(x_k,t)/\partial x_i^2 - \partial^2 p(x_k,t)/\partial t^2 = -Q(x_k,t), \tag{7}$$

all equivalent source terms (times c^2) now being represented by $Q(x_k,t)$. With these simplifications, one has a linear second order inhomogeneous partial differential equation, but with coefficients arbitrarily dependent on the space variables. Analytic solution of such equations is usually impracticable, being either impossible or very difficult, but numerical integration is now possible, by using methods such as Runge-Kutta and large computers (for long times (!)). This can be done also in cases when there are convective space derivative terms in p on the left-hand side, with specified spatially varying coefficients, as in the acoustics of sheared mean flows.

A reasonable variety of analytic methods are available for solution of equation (7) only when $c^2(x_k)$ is a constant, and similarly for convected forms of equation (7) only when the mean velocity is constant. This can always be arranged, of course, by a piecewise representation in which the domains of validity are taken small enough so that the mean temperatures and velocities, say, can be regarded as constants within them - but the domains may have to be taken so small that one is back, or nearly back, to the finite difference, lumped parameter forms of the equations so that there is no point in the continuum field, partial differential equation representation.

Despite these various difficulties, equation (7) with a constant $c \equiv c_o$ has a formal validity, and because of its linearity the effects of all components of the equivalent source distribution $Q(x_k,t)$ are additive. If the correct, physical values of $Q(x_k,t)$ are not used, the resulting approximate pressure field can be exactly corrected simply by adding to it the pressure field produced by the difference between the correct source and the approximate source that was used.

This feature of formal validity is of even more importance because of certain uniqueness and other properties of the solutions of the inhomogeneous scalar wave equation, with constant c (as assumed in what follows unless otherwise specified). The case of "constant" c does not exclude consideration of viscous and thermal attenuation effects on the waves, as these can be represented to a good approximation in most cases of interest by taking c to be of the form $c_o(1 + \tau_o \partial/\partial t)$, c_o being the lossless wave speed and the relaxation time τ_o being a linear function of the (constant) reference state viscosity, thermal conduction and other relaxation process coefficients. For simplicity, however, such losses are not included explicitly in what follows; in very many cases the attenuation per wavelength is very small and as far as the source power output is concerned radiation losses are dominant. Also, for simple harmonic motion (i.e., when the pressure spectral density is being considered, the lossy result can be obtained from the lossless one simply by replacing c by $c_o(1 + i\omega\tau_o)$, the Fourier transform convention here being

$$p(x_k,\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} p(x_k,t)\,e^{-i\omega t}dt, \qquad p(x_k,t) = \int_{-\infty}^{\infty} p(x_k,\omega)\,e^{i\omega t}d\omega. \tag{8}$$

Many of the important and useful properties of solutions of equation (7) with $c = c_o$, which can be written conveniently as

$$\left(\frac{\partial^2}{\partial x_k^2} - \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\right) p(x_k,t) = -q(x_k,t), \tag{9}$$

q being Q from equation (7) divided by c_o^2, can be obtained from the formal solution in terms of a Green function (i.e., a point-impulse response) and boundary values. The pressure $p(x_k,t)$ is assumed to satisfy equation (9) in a space-time domain D. The Green function $p_\delta(x_k,t)$ is then defined to satisfy

$$\left(\frac{\partial^2}{\partial x_k^2} - \frac{1}{c_o^2}\frac{\partial^2}{\partial t^2}\right) p_\delta(x_k,t;\, x_k',t') = -\delta(t - t')\,\delta(x_k - x_k') \tag{9a}$$

in D, with (x_k',t') also in D, $\delta(\,)$ being the Dirac delta function, and also to satisfy appropriate boundary and initial conditions (i.e., space time boundary conditions).

Then multiplying equation (9) by p_δ and equation (9a) by -p, adding the resulting equations and integrating over the domain D gives

$$p(x_k',t') = \int_D q\,p_\delta\,dD + \int_D \left(p_\delta\frac{\partial^2 p}{\partial x_k^2} - p\frac{\partial^2 p_\delta}{\partial x_k^2}\right)dD - \frac{1}{c_o^2}\int_D \left(p_\delta\frac{\partial^2 p}{\partial t^2} - p\frac{\partial^2 p_\delta}{\partial t^2}\right)dD. \tag{10}$$

The Green function can be taken to be

$$p_\delta = \delta(t' - t - |x_k - x_k'|/c_o)/4\pi|x_k - x_k'|, \tag{11}$$

which is a particular solution of equation (9a) (since $\delta(t - t') = \delta(t' - t)$, being an even function), plus any complementary function satisfying the homogeneous form of equation (9a), which in turn can be arbitrarily set equal to zero (since the initial and boundary conditions on the Green function, as an auxiliary function, can be arbitrarily specified, within reason). This form (11) for the Green function is not a customary one, such as can be found, with more detailed discussion of the physics, in Methods of Theoretical Physics, by P.M. Morse and H. Feshbach. The form (11) appears to violate causality, with $(t' - t)$ appearing instead of $(t' - t)$, but in fact this is perfectly valid mathematically because, as already mentioned, $\delta(t - t')$ is an even function. It is convenient to use the form here to shorten the

analysis without loss of mathematical rigour. With this choice of p_δ, and application of Green's theorem, equation (10) becomes

$$p(x_k',t') = \int_V \frac{q(x_k',\tau')}{4\pi|x_k' - x_k|}\,dx_k + \int_t \int_S \left(p_\delta \frac{\partial p}{\partial n} - p\frac{\partial p_\delta}{\partial n}\right) dS(x_k)\,dt - \frac{1}{c_o^2}\int_D \left(p_\delta \frac{\partial^2 p}{\partial t^2} - p\frac{\partial^2 p_\delta}{\partial t^2}\right) dD. \quad (12)$$

Here τ' is the retarded time $t' - |x_k' - x_k|/c_o$, V is the spatial volume implicit in the domain D and S is its boundary surface. The last integral in equation (12), over D explicitly, can be shown to represent the effects of initial conditions: that is to say, it represents disturbances from sources which have been turned on before the commencement of the time interval implicit in D, and which are hence completely independent (because of both causality and linearity) of the prescribed source distribution q. Because of the linearity, such effects, if there are any, can always be calculated separately, given knowledge of previous excitation, and added in at whatever stage in the calculations is convenient. The assumption, to be adopted here, that the integral is zero is equivalent to assuming that the system in D was quiescent before application of the specified source q. The first term in the integral explicitly over t and S in equation (12) self-evidently represents contributions from a surface source distribution of monopole type, of strength $\partial p/\partial n$ per unit area, as after integration over t it can be written simply as

$$\int_S \frac{1}{4\pi|x_k' - x_k|}\,\frac{\partial p}{\partial n}(x_k,\tau')\,dS(x_k).$$

The second term in this integral is similarly recognizable, upon replacing $\partial\delta/\partial n$ by its formal definition, in which n_j is the unit vector normal to the surface S,

$$\frac{\partial \delta}{\partial n} \equiv n_j \frac{\partial \delta}{\partial x_j} \equiv n_j \lim_{\varepsilon\to 0}\left\{\frac{\delta(t' - t - |x_k' - x_k - (\varepsilon/2)n_j|/c_o) - \delta(t' - t - |x_k' - x_k + (\varepsilon/2)n_j|/c_o)}{\varepsilon\, n_j}\right\},$$

as the contribution of a surface source distribution of dipole type, the dipole moment being in the direction normal to the surface and of magnitude $p(x_k,t)$. (From this formal definition this dipole term is readily seen to be the contribution of an impulsive monopole of strength $p(x_k,t)$ per unit area on one side of the surface together with one of equal and opposite strength, $-p(x_k,t)$, on the other side of the surface.)

Equation (12) thus shows that the pressure field $p(x_k',t')$ is the sum of elementary spherical wave contributions direct from each volume element of the source distribution (the first integral over V), and elementary spherical and dipole wave contributions from the boundary surface, each arriving at the observation point x_k' with the corresponding retarded times $\tau' = t' - |x_k' - x_k|/c_o$, to account for the time of travel $|x_k' - x_k|/c_o$ between the local source point x_k and the observation point x_k'. Dynamically, the monopole surface source distribution represents the rate of fluctuating injection of mass into the volume V and the dipole surface distribution the normal force on the material in V, both due to either the motion of the boundary or the reaction of the boundary, if motionless, on the material waves impinging on it. Obviously, in equation (12) the primed and unprimed variables can be interchanged and redefined, so that (x_k',t') are the space-time source co-ordinates and (x_k,t) those of the observation point, and it is notationally convenient to do this in what follows. For a point source, it is evident that the expressions in the two cases are identical, and this proves the "Principle of Reciprocity": that the pressure due to a point source at x_k' observed at x_k is the same as that which would be observed at x_k' if the point source were at x_k.

For a spatially unbounded domain D corresponding to negligible scattered or direct contributions from the surface S and with the non-zero source region confined to a finite region, the surface integral terms vanish, under assumption of the Sommerfeld radiation condition. It can then be shown that

$$\lim_{x_k'\to\infty} 4\pi|x_k|\,p(x_k,t)$$

is a four-dimensional Fourier transform of the source distribution $q(x_k,t)$ (see Chapter 1 of Noise and Acoustic Fatigue in Aeronautics, E.J. Richards and D.J. Mead, editors). Thus, because of the uniqueness of Fourier transform pairs, this far field radiated pressure, on the surface of a large sphere of radius $|x_k|$ wholly surrounding the non-zero source distribution region, has a unique one-to-one relationship with the source distribution, and from full knowledge of this far field the source distribution can be reconstructed. Babinet's principle can be regarded as a special case of this in which the source region is an infinitely thin illuminated aperture in a screen, and other cases where there is a surface source distribution rather than a volume source distribution can be similarly included as special cases. This result is an especially important one, and there can, and have been, misconceptions about it, as well as practical limitations in applying it. For example, an observer without access to the field over the whole sphere at large distances, and to a clock in the source region, cannot reconstruct the source: he could be fooled by "mirrors" or lack of knowledge of "source time"; also, he could be fooled by an applied mathematician using addition theorems to represent the field as due to a collection of physically impermissible but mathematically "correct" singular point monopoles, dipoles, quadrupoles, etc., at some arbitrary position and then misinterpreting his representation by saying that the actual source is thus not unique (!). Furthermore, there are obvious practical limitations because of the well-known half-wavelength resolving power: the structure of the source distribution with length scales smaller than a half-wavelength simply could not be determined in practice. Also, as mentioned, the theorem is not obtained even mathematically if singular source distributions are assumed to exist; such as point monopole, dipole, quadrupole, etc., distributions which are physically impossible, because they imply the existence of discontinuous "holes" in the continuum. Despite these limitations, this unique far field/source distribution relationship is the ultimate basis for development of "source-location" techniques, which are of increasing importance in practice.

Another uniqueness theorem which can be deduced from equation (12) by some further arguments is that the pressure (stress) field in the volume V is uniquely determined by the source distribution and the value of p, or $\partial p/\partial n$, or by the vanishing of a linear combination of p and $\partial p/\partial n$, say $A(S(x_k,t))p + \partial p/\partial n$, on the surface. Here $A(S(x_k))$ is in effect an admittance (which is the inverse of the specific acoustic impedance, which in turn is the ratio of p to the normal barycentric velocity at the surface). The function $A(S(x_k))$, as indicated by the notation, may vary arbitrarily from point to point, and in time, along the surface. This theorem can be established by taking $pA_p + \partial p/\partial n = 0$. Then

$$p_\delta \frac{\partial p}{\partial n} - p \frac{\partial p_\delta}{\partial n} = - p_\delta p A_p - p \frac{\partial p_\delta}{\partial n} . \qquad (13)$$

As has been mentioned, a boundary condition on p_δ, as an auxiliary function, can be prescribed arbitrarily, and also it can be shown that p as calculated from equation (12) is independent of this choice (because of the fact that p_δ as given in equation (11) is a particular solution of equation (9a), and any solution of the corresponding homogeneous equation can be added to it. Thus, if one takes p_δ as satisfying the same boundary condition as p, namely $Ap_\delta + \partial p_\delta/\partial n = 0$ (p_δ then being obviously the point impulse response of the material region of interest), the right-hand side of equation (13) evidently vanishes and hence so do the surface integral terms in equation (12). For this choice of the boundary condition on p_δ, therefore, only the volume integral term on the right-hand side of equation (12) is non-zero, and this then evidently includes both the waves directly emanating from the source region and those, previously emitted, which have been singly and multiply scattered at the boundaries before reaching the observer.

These interpretations of equation (12) are clearly in accord with experimental observations of the causality, propagation, and scattering properties of real wave phenomena.

They can be illustrated by working out the solution of the idealized, but by no means physically unrepresentative, case of a "point source" in an acoustic medium contained in a rectangular parallelepiped considered to have walls with constant pressure reflection coefficients, independent of angle of incidence. If the free field wave from the point source (at x_k') is $Lp(t - |x_k - x_k'|/c_o)/|x_k - x_k'|$, where L is a suitable constant with the dimensions of length, then the wave scattered from a plane surface, for which the vector from the source to the surface along the normal is h_k', is $RLp(t-|x_k - x_k' - 2h_k'|)/|x_k - x_k' - 2h_k'|$, the reflection coefficient R being a constant. If the rectangular parallelepiped walls are at $x_1 = (0,a)$, $x_2 = (0,b)$, and $x_3 = (0,d)$, and the six respective pressure wall reflection coefficients are denoted by $R(0,-,-)$, $R(a,-,-)$, $R(-,0,-)$, $R(-,b,-)$, etc., and the normal vectors similarly by $h_k'(0,-,-)$, $h_k'(a,-,-)$, etc., then the solution of equation (12) is, exactly,

$$p(x_k,t) = Lp(t - |x_k - k_x'|/c_o)/|x_k - x_k'| + R(0,-,-)Lp(t - |x_k - x_k' - 2h_k(0,-,-)|/c_o) + \ldots$$

$$+ R(0,-,-)R(-,0,-)Lp(t - |x_k - x_k' - 2h_k'(0,-,-) - 2h_k'(-,0,-)|/c_o) + \ldots + \text{etc.} \qquad (14)$$

The situation is clearly that of a three-dimensional "Versailles" hall of mirrors, but with imperfect mirrors, reflection from each mirror reducing the strength of the image source in the corresponding three-dimensional image space room by a factor $R(0,-,-)^\ell R(a,-,-)^n R(-,0,-)^m R(-,b,-)^q R(-,-,0)^r R(-,-,d)^s$, the integers ℓ, m, n, q, r, s corresponding to the numbers of times the wave, in travelling from the source to observation point, has been reflected from the $x_1 = (0,a)$, $x_2 = (0,b)$ and $x_3 = (0,d)$ walls. Evidently, in addition to the wave amplitude reductions on each reflection, there are reductions in proportion to the inverse of the distances the waves have travelled from the positions of their respective "image" source locations. Choice of the R's as ±1 corresponds to perfectly rigid and perfectly soft (stress release) walls, respectively, two extreme cases in neither of which is any energy transmitted to the exterior of the box, and values between -1 and 1 correspond to varying degrees of energy transmission, with all the R's zero being the free field condition in which only the direct wave, the first term of equation (14), remains. Thus in this simple example one has, in effect, a representation of all cases of energy loss(!). With order of magnitude guesses of the reflection coefficients, therefore, estimates of the relative importance of the direct radiation, and the wall-scattered radiation in all directions, can be easily made.

For real physical problems, neither dominance nor insignificance of one class of waves relative to the others usually occurs, unless by conscious design. When it is known to occur, obviously, processing and interpretation of data can be appropriately and appreciably simplified. When it does not, as is all too often the case, it must be kept firmly in mind that the observed pressure (stress) time-history is most probably composed of very many superposed stress wave components, perhaps each with a family resemblance to the direct signal (in the example the usual complication of dispersion upon each reflection was ignored - i.e., each reflected signal is usually spectrally distorted, not being just a reduced amplitude carbon copy of the incident signal, but almost certainly multiply overlapping in time. Determination of the principal paths along which energy arrives at the observation point then obviously becomes a difficult problem (!). The development of signal processing techniques, such as cepstral ones, by means of which contributions arriving at different times can be identified, is therefore of very great importance in acoustics and vibration problems.

The overall picture can be filled in with some reference to the relatively well-known techniques of modal analysis. Because the system is approximately equivalent, in finite difference terms, to a "piecewise" collection of finite degree of freedom systems, eigenfrequencies and eigenfunctions of the whole system are known to exist, and can be determined, albeit approximately, by suitable steady state, or other, testing methods or from operational models. For discussing the possibilities in this connection, it is appropriate to consider the pressure (stress) spectral density, as defined in equation (8). For the spectral density, equation (12) becomes simply

$$p(x_k,\omega) = \int_V \frac{q(x_k',\omega)\, e^{-i\omega|x_k - x_k'|/c_o}}{4\pi|x_k - x_k'|} + \int_S \left(\frac{e^{-i\omega|x_k - x_k'|/c_o}}{4\pi|x_k - x_k'|} \frac{\partial p}{\partial n'}(x_k',\omega) - \right.$$

$$- p(x_k',\omega) \frac{\partial}{\partial n'} \left(\frac{e^{-i\omega|x_k - x_k'|}}{4\pi|x_k - x_k'|} \right) dS(x_k') . \qquad (15)$$

With the eigenfrequencies and eigenfunctions known, the free space Green function spectral density

$$\frac{e^{-i\omega|x_k - x_k'|/c_o}}{4\pi|x_k - x_k'|} ,$$

or a Green function spectral density satisfying any other convenient boundary conditions, being of known analytical form, can be expressed as an expansion in the eigenfunctions (i.e., in effect, the orthonormal shape functions of the free vibrations of the system), or in any other convenient set of orthonormal functions suitable to the volume or surface under consideration. By performing the volume and surface integrations in equation (15), either numerically or analytically, as most convenient, $p(x_k,\omega)$ can then be obtained in the form of an expansion in these functions, $\psi_n(x_k;$ V or S), say, the "V or S" indicating that these functions depend on the volume or surface under consideration (it is important to note that these shape functions are purely geometrical in that they do not depend on the excitation frequency ω, or on any dynamical or material parameters such as the stress wave speed). This expansion will be of the form (the representation is simplified to emphasise the significant point)

$$p(x_k,\omega) = \sum_n a_n(\omega,c_o)\psi_n(x_k, \text{S or V}) , \qquad (16)$$

with the coefficients a_n carrying the frequency and material property information. Such representations have the very important property that, because of the orthonormality of the eigenfunctions (shape functions) ψ_n, the integral over the volume (surface) of the squared modulus of the pressure (stress) spectral density becomes simply

$$\int |p(x_k,\omega)|^2 dV(\text{or S}) = \sum_n |(a_n(\omega,c_o)|^2 . \qquad (17)$$

Since this volumetric (surface) average of $|p_n(x_k,\omega)|^2$ is representative of the average mean square pressure (stress) fluctuation over the volume (surface), representations such as equation (17), being a sum of positive definite quantities, depending only on frequency of excitation and material parameters, are obviously of great practical use (nothing is easier than adding up a series of positive numbers, provided it converges quickly enough !).

As well as global mean square stresses, mean energy flux due to fluctuating motion can be expressed in terms similar to equation (17). What is not evident in such expressions is that which is evident, in all its complications, from expressions such as equations (12) and (14), which show clearly the intricacies of interference (i.e., superposition of signals from different space-time origins) that inevitably complicate most wave phenomena. Unless a wave is present it cannot possibly carry energy, whether or not it is subject to interference from other waves.

An important, and little-known result, applicable to pressure (stress) spectral density is that a mean value theorem applies to the homogeneous scalar Helmholtz equation $(\partial^2 p/\partial x_i^2) + k^2 p = 0$, k and p real, as well as to the Laplace equation $\partial^2\phi/\partial x_i^2 = 0$. From the Laplace equation, the mean value theorem is that at a point x_k, surrounded by a surface S, $\phi(x_k)$ is equal to its average value over the surface S; i.e.,

$$\phi(x_k) = \frac{1}{S} \int_S \phi(x_k') dS(x_k') .$$

For the scalar Helmholtz equation, the result is rather different (it appears to be due to Pöckels (1886) and has been largely ignored since): it is that for any sphere of radius $n\lambda/2$, where λ, the wavelength, is equal to $k/2\pi$ and n is any positive integer, the value of p averaged over the volume interior to the sphere is zero. This means that p has at least one null surface passing through this sphere, and, thus, broadly speaking, that any pressure (stress) standing wave field of frequency ω has negative minima (or positive maxima) separated by, on spatial average, a wavelength, which is not surprising in view of known solutions like sin (or cos) k_x, $(\sin kr)/r$, etc.

However, a deduction from this result is somewhat more surprising: it is that the amplitude of any harmonic standing wave field, of frequency ω, being expressible as $|p| = \sqrt{p_R^2 + p_I^2}$, where both the real and imaginary parts p_R and p_I satisfy the mean value theorem, is the square root of the sum of the squares of only two such spatially oscillating functions (as well as of an infinity of particular spatially oscillating functions, of all wavelengths, as implied by eigenfunction representation). By working back from this it can be proved that a plane progressive harmonic wave is the only pressure (stress) wave having a spatially constant amplitude.

Conclusions and recommended guidelines for the future are as follows, in an order which is not that followed in the preceding discussion.

(i) Because of their well-known convenient mathematical properties, discrete or continuous orthogonal representations of stress wave fields (eigenfunction series, or Fourier frequency or wavenumber transforms) are at present the most reliable for analytical and/or computational purposes.

(ii) When wave signal information is required, rather than wave amplitude averages, however, the representations cited in conclusion (i) are very often inconvenient, or not at all tractable; for such situations new methods, such as cepstral analysis, need to be further developed to

provide results more realistically interpretable in terms of the complicated space-time propagation paths and interference patterns characteristic of real wave motion when single and multiple scattering is present.

(iii) In the development of improved techniques for analysis of wave field data, account should be taken, and use could well be made, of certain fundamental uniqueness and other properties of wave fields.

VIBRATION OF STRUCTURES EXCITED ACOUSTICALLY

B.L. Clarkson
Professor of Vibration Studies
Institute of Sound & Vibration Research
University of Southampton
Southampton SO9 5NH
England

SUMMARY

When a sound wave impinges on a structure, vibrations are set up in the structure and vibration and sound energy are transmitted. These two lectures develop the analytical methods available to estimate the response of structures on the assumption that the structure itself does not change the properties of the incident sound field by any feed back or flutter type mechanism. The normal mode method is used for the majority of the work to estimate the response to any sound field. Special cases considered are jet noise excitation, turbulent boundary layer pressure excitation and a reverberant field excitation. The alternative travelling wave method is also introduced.

LIST OF SYMBOLS

A	surface area
	constant
a	speed of travel of pressure wave
b	width of plate
C	generalised damping
c	speed of convection over surface of structure
D	flexural rigidity
E	Young's modulus
e	exponential
F	force
h	plate thickness
i	imaginary symbol
j	joint acceptance function
K	generalised stiffness
k	flexural wave number
L	total load on structure
ℓ	length of beam or plate
M	generalised mass
	bending moment
m	mass per unit area
P	amplitude of harmonic pressure
	Fourier transform of p
p	pressure
q	generalised co-ordinate
R	correlation function
r,s	mode numbers
S	spectral density function
t	time
w	displacement
X	real part of frequency function
x	position on structure
Y	imaginary part of frequency function
	wave receptance function
Z	impedance
α	receptance (frequency response function for displacement)
δ	Dirac delta function
ε	phase angle
η	loss factor
θ	angle of incidence
λ	wavelength
λ_t	trace wavelength
μ	component of receptance which is dependent on mode shape
ν	Poisson's ratio
Π	power flow
ρ	correlation coefficient
	density
τ	time delay
$\phi(x)$	displacement mode shape
ϕ	phase angle
ω	frequency rads/sec
$-$	indicates vector quantity
κ	radius of gyration of section

INTRODUCTION

An acoustic wave impinging on a structure produces a time varying pressure load on the structure which sets up vibrations of the surface plates and leads to transmission of vibration to internal structures and equipment. There are two major analytical methods available for determining the structural response. The first and perhaps most widespread in use is the normal mode method and its practical implementation through the finite element representation of the structure. The second method considers the wave motion which is set up in the structure by a surface force. For certain types of structure this produces a very elegant solution. The major part of these notes is devoted to the normal mode method.

Several models for the acoustic loading have been developed and the appropriate one can be chosen for the particular problem under consideration. The simplest representation is a harmonic plane wave impinging on the structure at a fixed angle. This can be extended to include many wavelengths added together in a random manner. Alternatively the angle of incidence can be assumed to be random as in a diffuse field. In another class of problems the pressure field can be represented by a pressure pattern travelling over the structure. The simplest representation is a 'frozen' convected pattern where there are no changes in the convected frame of reference. A more realistic representation allows for changes by assuming decay in the statistical properties. Finally a completely random field in space and time can be assumed.

The analytical methods will be outlined briefly and then the results for acoustic excitation will be derived. Using the normal mode method results will be obtained for all the loading cases mentioned above. Then the travelling wave result for the travelling wave excitation will be derived using the wave propagation method.

THE NORMAL MODE METHOD

In the absence of damping it is assumed that the structure can vibrate freely in a series of normal modes:

$$\phi_1(x) \quad \phi_2(x) \quad \ldots\ldots\ldots \quad \phi_r(x) \ .$$

During any arbitrary vibration the displacement of any point x on the structure can be represented by

$$w(x,t) = \Sigma_r \phi_r(x)\, q_r(t) \qquad (1)$$

$q_r(t)$ is the generalised co-ordinate of the r^{th} mode and $\phi_r(x)$ is the mode shape. This gives a separable solution in x and t.

When the system is subjected to a time dependent load, the displacement $q_r(t)$ of any one of the normal modes of vibration is governed by the same type of equation as the simple spring-mass-damper oscillator. From considerations of the energy of the system we can write the Lagrange equations as

$$M_r \ddot{q}_r + C_r \dot{q}_r + K_r q_r = F_r(t) \qquad (2)$$

where

$M_r = \int_A m \phi_r^2(x)\, dA$ — generalised mass

$C_r = \int_A p_{dr}(x)\, \phi_r^2(x)\, dA$ — generalised damping

$K_r = \omega_r^2 M_r$ — generalised stiffness

$F_r(t) = \int_A p(xt)\, \phi_r(x)\, dA$ — generalised force

m = mass per unit area

P_{dr} = local viscous damping pressure per unit velocity.

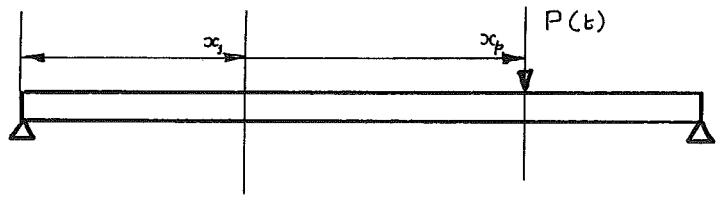
Cross terms do not appear in equations for r^{th} mode due to orthogonality of normal modes. Actually there will be some cross coupling of the modes due to work done by damping forces of one mode in moving through displacements of another mode. For small damping this type of coupling can be neglected. For high damping it is possible to use one of the techniques suggested by Fraeys de Veubeke, Mead and others, where a series of uncoupled damped modes are derived.

Alternatively we can rewrite the Lagrange equation in terms of hysteretic damping as:

$$M_r \ddot{q}_r + M_r \omega_r^2 (1 + i\eta_r)\, q_r = F_r(t) \qquad (3)$$

where η_r = the loss factor in the r^{th} mode.

As an example of the use of this technique we can calculate the response of a beam to a point harmonic force $P(t) = P_o e^{i\omega t}$



The displacement of any point x_1 is

$$w(x_1 t) = \sum_r \phi_r(x_1)\, q_r(t)$$

The generalised force is

$$F_r(t) = P_o e^{i\omega t}\, \phi_r(x_p)$$

Then the equations of motion can be written

$$M_r \ddot{q}_r + \omega_r^2 (1 + i\eta_r) M_r q_r = \phi_r(x_p)\, P_o e^{i\omega t} \qquad r = 1,2 \ldots$$

Solving we get

$$q_r(t) = \frac{\phi_r(x_p)\, P_o e^{i\omega t}}{M_r(\omega_r^2 - \omega^2 + i\eta_r \omega_r^2)}$$

Therefore the displacement at point x_1 is

$$w(x_1 t) = \sum_r \frac{\phi_r(x_1)\, \phi_r(x_p)}{M_r(\omega_r^2 - \omega^2 + i\eta_r \omega_r^2)}\, P_o e^{i\omega t}$$

which can be written as $\underline{\underline{w(x_1 t) = \alpha(x_1 x_p \omega)\, P_o e^{i\omega t} = \alpha(x_1 x_p \omega)\, P(t)}}$

$\alpha(x_1 x_p \omega)$ is the frequency response function of the structure. In our usage it gives the displacement response of the structure at position x_1 to a unit magnitude harmonic force at x_p. It is known as the receptance function and can be written as.

$$\alpha(x_1 x_p \omega) = \sum_r \frac{\phi_r(x_1)\, \phi_r(x_p)}{M_r} \cdot \frac{1}{(\omega_r^2 - \omega^2 + i\eta_r \omega_r^2)}$$

$$= \sum_r \frac{\phi_r(x_1)\, \phi_r(x_p)}{M_r} \cdot \frac{(\omega_r^2 - \omega^2) - i\eta_r \omega_r^2}{(\omega_r^2 - \omega^2)^2 + \eta_r^2 \omega_r^4} \qquad (4)$$

which can conveniently be split up into real and imaginary parts

$$\alpha(x_1 x_p \omega) = \sum_r \frac{\phi_r(x_1)\, \phi_r(x_p)}{M_r} (X_r - iY_r)$$

or $\underline{\underline{\alpha(x_1 x_p \omega) = \sum_r \mu_r (X_r - iY_r)}} \qquad (5)$

where

$$X_r = \frac{\omega_r^2 - \omega^2}{(\omega_r^2 - \omega^2)^2 + \eta_r^2 \omega_r^4}$$

$$Y_r = \frac{+\eta_r \omega_r^2}{(\omega_r^2 - \omega^2)^2 + \eta_r^2 \omega_r^4}$$

$$\mu_r = \frac{\phi_r(x_1)\, \phi_r(x_p)}{M_r}$$

Thus it can be seen that the receptance is made up of two types of terms. The μ_r is dependent on the mode shapes of the structure and the X,Y terms are dependent on the frequency and loss factor. The frequency terms give the well known resonance peak and phase shift characteristics at the undamped natural frequency of the r^{th} mode. Either or both of the $\phi_r(x)$ terms could be zero depending on the positions of the displacement measuring point and the excitation point.

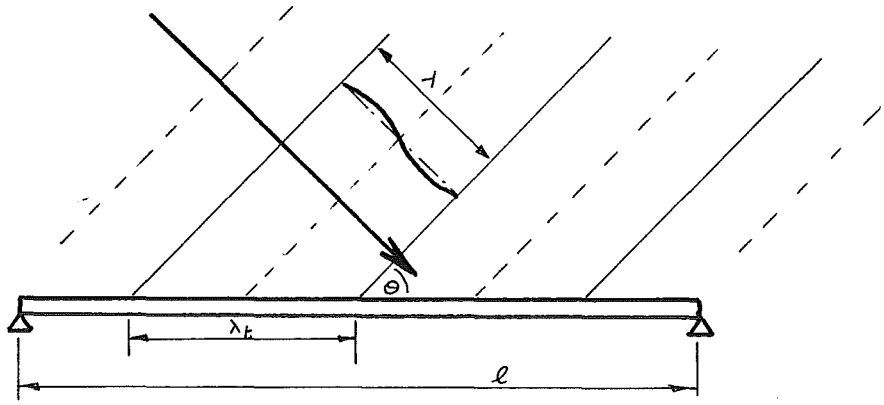
THE RESPONSE OF A PLATE TO A PLANE WAVE

The idea of coupling of the pressure field with the structural deformation pattern can be introduced best by the consideration of the response of a plate to a travelling field of harmonic waves. Suppose such a field impinges upon a plate, the wavefronts making an angle θ with the plate surface. Consider only the case when the pressure wave motion has no component of velocity across the width of the plate, i.e. the wave travels in the direction of the LENGTH of the plate.

The wavefronts make an intercept of λ_t on the plate. We call λ_t the 'trace wavelength', and the speed at which the wavefront moves along the plate the 'trace velocity' (a_t). Clearly, $\lambda_t = \lambda \operatorname{cosec} \theta$, and $a_t = a \operatorname{cosec} \theta$ (a = speed of propagation).

The instantaneous pressure at any point along the plate is

$$p(x,t) = p_o \cos\left(\omega t - \frac{2\pi x}{\lambda_t} + \phi\right). \qquad (6)$$



Assume that the plate (dimensions ℓ x b) is simply supported along its edges. The effect of the number of modal wavelengths along the propagating direction can be found by considering the (r,1) mode:

$$\phi_r\,(xy) \;=\; \sin r\,\frac{\pi x}{\ell}\;\sin\frac{\pi y}{b}$$

The generalised force is given by:

$$F_r(t) \;=\; \int_0^b\int_0^\ell p_o\,\cos(\omega t - \frac{2\pi x}{\lambda_t})\,\sin\frac{r\pi x}{\ell}\,\sin\frac{\pi y}{b}\,dx\,dy$$

$$=\; p_o\,\frac{2b}{\pi}\int_0^\ell \cos\,(\omega t - \frac{2\pi x}{\lambda_t})\,\sin\frac{r\pi x}{\ell}\,dx\;.$$

Evaluating the integral, and replacing ℓ/r by $\lambda_m/2$ (the half-wavelength of the mode of vibration) we find:

$$F_r(t) \;=\; p_o\ell b\,\frac{4}{\pi^2}\,\{\,\frac{\cos(\frac{\pi}{2}\,r\lambda_m/\lambda_t)}{r\{(\lambda_m/\lambda_t)^2 - 1\}}\,\}\;\cos\,(\omega t + \varepsilon) \qquad (r\ \text{odd}) \qquad (7)$$

$$\text{or} \qquad =\; p_o\ell b\,\frac{4}{\pi^2}\,\{\,\frac{\sin(\frac{\pi}{2}\,r\lambda_m/\lambda_t)}{r\{(\lambda_m/\lambda_t)^2 - 1\}}\;\cos\,(\omega t + \varepsilon) \qquad (r\ \text{even})$$

The Joint Acceptance Function

Equation (7) may be written in the form:

$$F_r(t) \;=\; p_o\ell b j_r\,\cos\,(\omega t + \varepsilon)\;. \qquad (8)$$

The total force acting on the plate at any instant when the pressure is distributed uniformly over the plate is $p_o\ell b\,\cos\,(\omega t + \varepsilon)$. j_r is always less than one, and is a factor which describes the proportion of this force which a particular mode of distortion can 'accept' and convert into the corresponding generalised force. Since it is a function of both the modal wavelength, λ_m, and the trace wavelength, λ_t, we may call it the JOINT ACCEPTANCE of the mode and pressure field. Figure 1 shows how j_r varies with both r and λ_m/λ_t. Apart from the curve for r=1, it is characteristic of each curve that its maximum value occurs at $(\lambda_m/\lambda_t) = 1$, i.e. when the modal wavelength is the same as the trace wavelength. The response of the plate in any of its modes to this pressure field can be obtained as follows:

The generalised response function of the plate can be written as:

$$q_r(t) \;=\; p_o\ell b\,|\alpha_r|\;j_r\,\cos\,(\omega t + \varepsilon - \phi_r)\;. \qquad (9)$$

It is now clear that a large response in this mode will occur if, at the resonant frequency ω_r (when $|\alpha_r|$ is a maximum) j_r is at the maximum value corresponding to $\lambda_m/\lambda_t = 1$. A dual coincidence then occurs, the sound field frequency being equal to the resonant frequency and the trace wavelength being equal to the modal wavelength. For this reason the term COINCIDENCE EFFECT is often used to describe this particular phenomenon.

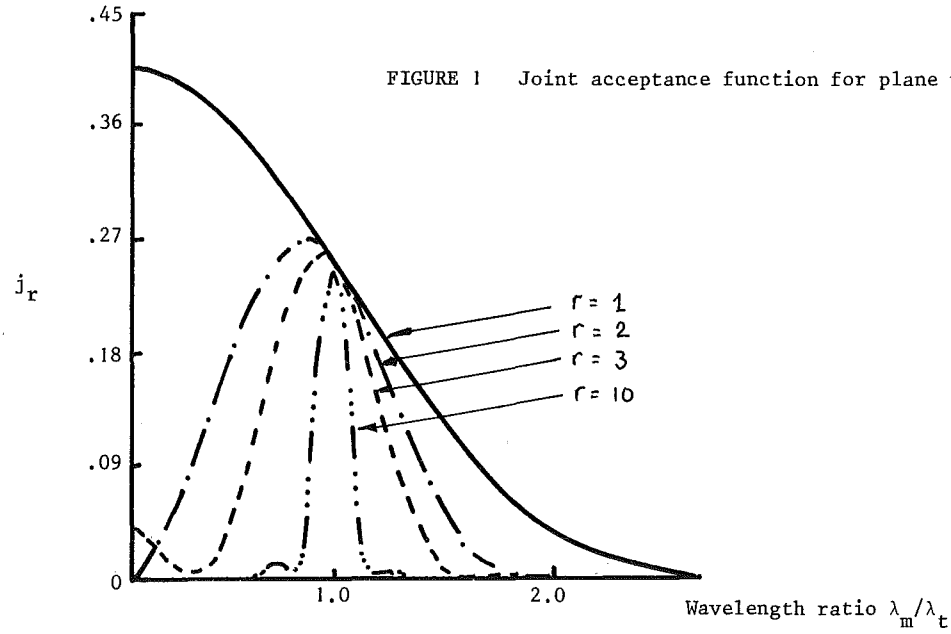
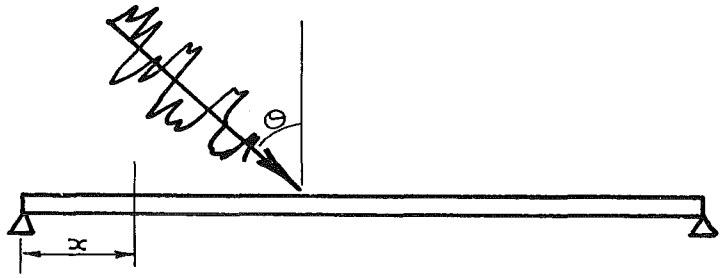


FIGURE 1 Joint acceptance function for plane waves

THE GENERALISED FORCE PRODUCED ON A BEAM BY RANDOM PLANE WAVES

The next extension of the simple case described in the preceding section is to include in the travelling wave a combination of wavelengths such that the pressure at any one point is random but the whole pattern moves in the direction θ with velocity a. This might be used to represent waves in the deep ocean where the wave motion is propagating in the direction of the surface wind or as the simplest approximation for boundary layer or pipe flow excitation.



The pressure at a point can be written in the Fourier integral form:

$$p(t) = \int_{-\infty}^{\infty} P(i\omega)\, e^{i\omega t}\, d\omega$$

Physically, this is equivalent to representing the oncoming wave as the sum of an infinite number of waves of different wavelength.

The pressure distribution on the beam is then given by:

$$p(t,x) = \int_{-\infty}^{\infty} P(i\omega)\, e^{i\omega(t - \frac{x}{a}\sin\theta)}\, d\omega \qquad (10)$$

The total load on the beam is:

$$F(t) = \int_0^{\ell} p(t,x)\, dx = \int_0^{\ell}\int_{-\infty}^{\infty} P(i\omega)\, e^{i\omega(t - \frac{x}{a}\sin\theta)}\, d\omega\, dx.$$

writing $\omega \frac{x}{a} \sin\theta = \frac{2\pi x}{\lambda_t}$ and performing the x integration gives:

$$F(t) = \int_{-\infty}^{\infty} \{ P(i\omega) \frac{i\lambda t}{2\pi} (e^{\frac{-i2\pi\ell}{\lambda t}} - 1) \}\, e^{i\omega t}\, d\omega.$$

The expression in { } brackets is the Fourier transform $F(i\omega)$ of $F(t)$ and therefore the spectral density of $F(t)$ is

$$S_F(\omega) = \lim_{T\to\infty} \frac{\pi}{T} |F(i\omega)|^2 = \lim_{T\to\infty} \frac{\pi}{T} |P(i\omega)|^2 \left\{ \frac{\sin \frac{\pi\ell}{\lambda_t}}{\frac{\pi\ell}{\lambda_t}} \right\}^2 \ell^2$$

$$= S_p(\omega)\ell^2 \left\{ \frac{\sin \frac{\pi\ell}{\lambda_t}}{\frac{\pi\ell}{\lambda_t}} \right\}^2 \qquad (11)$$

$$= S_p(\omega)\, \ell^2\, j^2$$

Figure 2 shows the variation of j^2 with wavelength ratio. As the spectral density is written in terms of squares of force etc. the joint acceptance function appears as j^2. If the pressure is exactly in phase over the whole of the beam the spectral density of load is $S_p(\omega)\ell^2$. The joint acceptance function j^2 gives a measure of the coupling of a particular pressure field with the beam. This equation (11) demonstrates a very important point, viz. that by virtue of its length the beam acts as a 'wavelength filter'. Whenever the frequency ω is such that $\pi\ell/\lambda_t$ $(= \omega\ell \sin\theta/2a)$ is an integer, then the corresponding spectral density of the total force is zero. The principal part of the generalised force comes from those components, the wavelengths of which are long compared with the length of the beam. That this must be so can be seen from simple physical considerations; a wave component whose trace wavelength is exactly equal to the span of the beam provides a positive pressure over one half of the span and a negative pressure over the other half. The corresponding net force is therefore zero.

A similar calculation may be carried out to determine the spectral density of the generalised force corresponding to the mode of displacement $\phi_r(x) = \sin(\pi rx/\ell)$. The generalised force, $F_r(t)$, is given by:

$$F_r(t) = \int_0^\ell p(t,x) \sin \frac{\pi rx}{\ell}\, dx.$$

Substituting for $p(t,x)$ from equation (10), and carrying out the integration in the same way as above, we find that the spectral density of $F_r(t)$ is

$$S_F(\omega) = S_p(\omega)\ell^2 \frac{4}{\pi^2} \left\{ \frac{\sin\left[(\pi/2)r\lambda_m/\lambda_t + r\pi/2\right]}{r\left[(\lambda_m/\lambda_\ell)^2 - 1\right]} \right\}^2 \qquad (12)$$

in which $\lambda_m/2 = \ell/r$.

If we had considered a plate, vibrating in the mode $\sin(\pi rx/\ell) \sin(\pi y/b)$, this expression would have become:

$$S_F(\omega) = S_p(\omega)(\ell b)^2 \frac{16}{\pi^4} \left\{ \frac{\sin\left[(\pi/2)r\lambda_m/\lambda_t + r\pi/2\right]}{r\left[(\lambda_m/\lambda_t)^2 - 1\right]} \right\}^2 \qquad (13)$$

or

$$S_F(\omega) = S_p(\omega)(\ell b)^2\, j_r^2 \qquad (14)$$

The wavelength 'filter' characteristics of the plate are again seen by the manner in which j_r (and hence j_r^2) varies with ℓ/λ_t. Again we have the spectral density of the generalised force given as the product of the power spectral density of the (homogeneous) pressure, the square of the plate area and the square of the joint acceptance.

THE GENERALISED FORCE CORRESPONDING TO AN ENTIRELY RANDOM, HOMOGENEOUS PRESSURE FIELD

We now remove the restriction of the former paragraph that the sound waves must be plane and moving down the length of the plate. Further, we shall consider a general mode shape, $\phi_r(x,y)$ which is not necessarily a sinusoidal mode.

The generalised force is

$$F_r(t) = \int_A \phi_r(x,y).\, p(x,y;t)\, dA$$

where the integration is now over the surface. The Fourier integral transform of this is:

$$F_r(i\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_A \phi_r(x,y).\, p(x,y;t)\, e^{i\omega t}\, dA\, dt$$

$$= \int_A P(i\omega)\, \phi_r(x,y)\, dA$$

in which $P(i\omega)$ is the Fourier transform of $p(x,y;t)$. The spectral density of the generalised force, $S_{F_r}(\omega)$ is given by:

$$S_{F_r}(\omega) = \lim_{t_o \to \infty} \frac{\pi}{t_o} |F_r(i\omega)|^2 = \lim_{t_o \to \infty} \frac{1}{t_o} \frac{1}{4\pi} \int_A \int_A \phi_r(x_1 y_1)\, \phi_r(x_2 y_2) \int_{-t_o}^{+t_o} \int_{-t_o}^{+t_o} p(x_1 y_1 t_1)\, p(x_2 y_2 t_2)\, e^{i\omega(t_1 - t_2)}\, dt_1\, dt_2\, dA_1\, dA_2 \quad . \tag{15}$$

Now denote the point $x_1 y_1$ by $\bar{x}_1$, $x_2 y_2$ by $\bar{x}_2$, and $t_2 - t_1$ by τ.

Equation (15) becomes

$$S_{F_r}(\omega) = \int_A \int_A \phi_r(\bar{x}_1)\, \phi_r(\bar{x}_2) \lim_{t_o \to \infty} \frac{1}{t_o} \frac{1}{4\pi} \int_{-t_o}^{+t_o} \int_{-t_o}^{+t_o} p(\bar{x}_1, t_1)\, p(\bar{x}_2, t_1 + \tau)\, e^{-i\omega t}\, dt_1\, d\tau\, dA_1\, dA_2 \tag{16}$$

But

$$\lim_{t_o \to \infty} \frac{1}{2t_o} \int_{-t_o}^{+t_o} p(\bar{x}_1, t_1)\, p(\bar{x}_2, t_1 + \tau)\, dt_1 = R_p(\bar{x}_1, \bar{x}_2, \tau) \tag{17}$$

which is the cross-correlation function of the pressures at $\bar{x}_1$ and $\bar{x}_2$ and is sometimes known as the 'space -time correlation function'.

The limit of the time integral contained within the right-hand side of equation (16) now becomes:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} R_p(\bar{x}_1, \bar{x}_2, \tau)\, e^{-i\omega\tau}\, d\tau = S_p(\bar{x}_1, \bar{x}_2, \omega) \tag{18}$$

which is the 'cross spectral density' of the pressures at $\bar{x}_1$ and $\bar{x}_2$. Since $R_p(\bar{x}_1, \bar{x}_2,)$ is not an even function of τ, the cross spectral density will be a complex quantity.

$$S_{F_r}(\omega) = \int_A \phi_r(\bar{x}_1) \int_A \phi_r(\bar{x}_2)\, S_p(\bar{x}_1, \bar{x}_2, \omega)\, dA_2\, dA_1 \tag{19}$$

When this integral is examined in detail, it will be found that the imaginary component vanishes identically (as it must, since $S_{F_r}(\omega)$ is essentially real). For the purposes of evaluating the integral of equation (19), we therefore need only the real part of $S_p(\bar{x}_1, \bar{x}_2, \omega)$, i.e.

$$\frac{1}{2\pi} \int_{\infty}^{\infty} R_p(\bar{x}_1, \bar{x}_2, \tau) \cos \omega\tau\, d\tau$$

This is identical to the 'narrow-band' space correlation of the pressures at the two points, $\bar{x}_1$ and $\bar{x}_2$.

The cross spectral density may be non-dimensionalised by dividing it by the direct spectral density, $S_p(\omega)$ of the (homogenous) pressure field. Equation (19) becomes:

$$S_{F_r}(\omega) = S_p(\omega) \int_A \phi_r(\bar{x}_1) \int_A \phi_r(\bar{x}_2)\, \rho(\bar{x}_1, \bar{x}_2, \omega)\, dA_2\, dA_1 \tag{20}$$

$\rho(\bar{x}_1, \bar{x}_2 \omega)$ has the form of a correlation coefficient, and being a function of ω, may be regarded as the spectrum of the correlation coefficient.

The inner integral

$$\int_A \phi_r(\bar{x}_2)\, \rho(\bar{x}_1, \bar{x}_2, \omega)\, dA_2$$

has the dimensions of an area, and may be regarded as a correlation area. In a homogeneous field $\rho(\bar{x}_1, \bar{x}_2, \omega)$ is a function of the separation ξ, of the points x_1 and x_2 and not of x_1 alone. The inner integral is therefore the area under the product of two curves, as in Figure 3.

At high frequencies the distance to the first zero crossing of $\rho(\bar{x}_1, \bar{x}_2, \omega)$ is small compared with the wavelength of $\phi(x)$, and the integral is approximately equal to

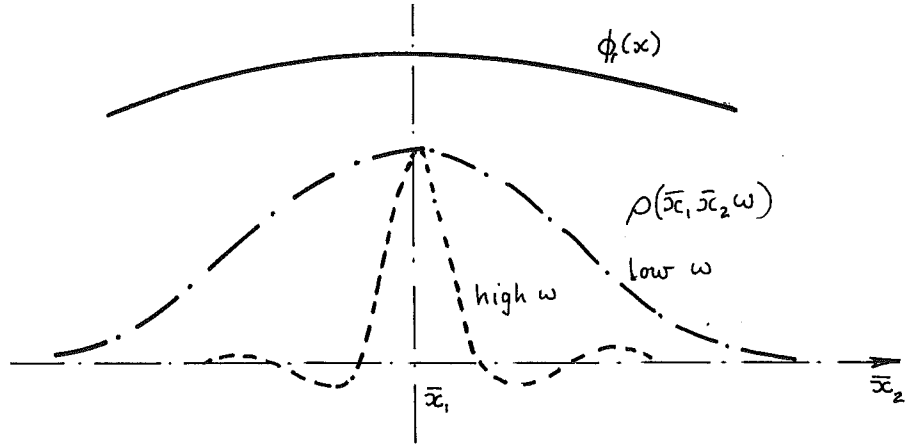
$$\phi_r(\bar{x}_1) \int_A \rho(\bar{x}_1, \bar{x}_2, \omega)\, dA_2$$

which may be written in the form

$$\phi_r(\bar{x}_1)\, \lambda^2(\omega) \quad .$$

$\lambda^2(\omega)$ is the correlation area at the frequency ω and is independent of either $\bar{x}_1$ or $\bar{x}_2$.

FIGURE 3 Illustration of coupling between mode shape and correlation functions



Equation (20) now becomes

$$S_{F_r}(\omega) \simeq S_p(\omega)\ \lambda^2(\omega) \int_A \phi_r^2(\bar{x})\ dA \qquad (21)$$

which is the approximate expression for the spectral density of the generalised force at high frequencies.

At low frequencies when $\rho(\bar{x}_1, \bar{x}_2, \omega)$ is nearly constant ($\simeq 1$) all over the surface, equation (20) becomes

$$S_{F_r}(\omega) \simeq S_p(\omega) \left[\int_A \phi_r(\bar{x})\ dA\right]^2 \qquad (22)$$

Now when ϕ_r represents a sinusoidal mode of r half-waves over the surface area, the integral in equation (22) will be zero when r is even, and when r is odd it will get increasingly smaller as r increases. The spectral density at low frequencies may therefore have zero values, and will be very small for the high order modes. The integral in equation (21) can never be zero as the integrand is squared. For sinusoidal modes, the integral is independent of r, and is equal to one-quarter of the plate area. At high frequencies therefore, the spectral density of the generalised force in all the modes is proportional to the magnitude of the correlation area.

Finally we write equation (20) in the form:

$$S_{F_r}(\omega) = S_p(\omega)\ A^2 \frac{\int_A \int_A \phi_r(\bar{x}_1)\ \phi_r(\bar{x}_2)\ \rho(\bar{x}_1, \bar{x}_2, \omega)\ dA_1\ dA_2}{A^2}$$

$$= S_p(\omega)\ A^2\ j_{rr}^2 \qquad (23)$$

which defines the general expression for the 'direct' joint acceptance, j_{rr}.

CALCULATION OF THE TOTAL RESPONSE

We have now seen how to calculate the generalised force in certain special cases, and also how to calculate the response in just one mode. At first sight one might think that the total mean square response could be obtained by summing the mean square responses of the individual contributory modes, but this may only be done if the responses in the various modes are statistically independent. This is not so here, for a particular frequency component of the pressure fluctuations can excite vibrations in several modes, albeit by small amounts in some of them. The corresponding frequency components in the fluctuating response of these modes will be correlated to a certain extent.

The amplitude of motion in the r'th mode is given by the generalised co-ordinate:

$$q_r(t) = \frac{F_r(t)}{M_r(\omega_r^2 - \omega^2 + i\eta_r \omega^2)} \qquad \text{where } F_r(t) \text{ is the generalised force.}$$

$$= \frac{F_r(t)}{M_r} (X_r - iY_r) \qquad \text{using the notation of (5).}$$

The generalised force can be written in its Fourier transform form to give:

$$q_r(t) = \int_{-\infty}^{\infty} \frac{F_r(i\omega)}{M_r} (X_r - iY_r)\ e^{i\omega t}\ d\omega$$

The displacement in the r'th mode at any point x on the vibrating structure is

$$w_r(x,t) = q_r(t)\,\phi_r(x) = \int_{-\infty}^{\infty} F_r(i\omega)\,\frac{\phi_r(x)}{M_r}\,(X_r - iY_r)\,e^{i\omega t}\,d\omega$$

having the Fourier transform $F_r(i\omega)\,\frac{\phi_r(x)}{M_r}\,(X_r - iY_r)$.

The corresponding spectral density is therefore given by:

$$S_w(\omega) = \lim_{t_o\to\infty} \frac{\pi}{t_o} \sum_{r=1}^{\infty} F_r(i\omega)\,\frac{\phi_r(x)}{M_r}\,(X_r - iY_r) \sum_{r=1}^{\infty} F_r^*(i\omega)\,\frac{\phi_r(x)}{M_r}\,(X_r + iY_r)$$

On multiplying this out and proceeding to the limit of $t_o \to \infty$ we find

$$S_w(\omega) = \sum_{r=1}^{\infty} \frac{\phi_r^2(x)}{M_r^2}\,(X_r^2 + Y_r^2) \int_A \phi_r(\bar{x}_1) \int_A \phi_r(\bar{x}_2)\,S_p(\bar{x}_1\bar{x}_2\omega)\,dA_2\,dA_1$$

$$+ \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \frac{\phi_r(x)\phi_s(x)}{M_r M_s}\,(X_r - iY_r)\,(X_s + iY_s) \int_A \phi_r(\bar{x}_1) \int_A \phi_s(\bar{x}_2)\,S_p(\bar{x}_1\bar{x}_2\omega)\,dA_2\,dA_1 \qquad (24)$$

$r \neq s$

The first series gives the sum of the spectra of the responses in the individual modes. The second (double) series is the 'correction' term which is necessary due to the correlation between the responses in different modes, and is said to be the response due to 'modal frequency overlap'.

It can be seen that both the real and imaginary parts of $S_p(\bar{x}_1,\bar{x}_2,\omega)$ contribute to the second series. The final value of this is, of course, entirely real. This is in contrast to the first series, to which only the real part contributes, as was shown in the last section.

The double integral in the first series is the spectral density of the r'th generalized force, and may be written in the form of equation (23), involving the joint acceptance. It then becomes

$$A^2\,S_p(\omega) \sum_{r=1}^{\infty} \frac{\phi_r^2(x)}{M_r^2}\,(X_r^2 + Y_r^2)\,j_{rr}^2$$

The second series may be manipulated in a similar way to yield

$$A^2\,S_p(\omega) \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \frac{\phi_r(x)\phi_s(x)}{M_r M_s}\,(X_r - iY_r)(X_s + iY_s)\,j_{rs}^2$$

Equation (24) can then be written in a simpler form as

$$S_w(\bar{x}_1\omega) = S_p(\omega)\,A^2 \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \frac{\phi_r(x_1)\,\phi_s(x_1)}{M_r M_s}\,(X_r - iY_r)(X_s + iY_s)\,j_{rs}^2 \qquad (25)$$

where joint acceptance term j_{rs}^2 is given by

$$j_{rs}^2 = \frac{1}{S_p(\omega)A^2} \int_A \int_A \phi_r(\bar{x}_A)\,\phi_s(\bar{x}_B)\,S_p(\bar{x}_A\,\bar{x}_B\omega)\,d\bar{x}_A\,d\bar{x}_B \qquad (26)$$

This has been normalised by $S_p(\omega)\,A^2$.

Alternatively equation (25) can be written in terms of the receptance α

we note that $$\sum_{r=1}^{\infty} \frac{\phi_r(\bar{x}_1)}{M_r}\,(X_r - iY_r)\,\phi_r(\bar{x}_A) = \alpha_{IA}$$

therefore we can write:

$$S_w(\bar{x}_1\omega) = \int_A \int_A \alpha(\bar{x}_1\bar{x}_A\omega)\;\alpha^*(\bar{x}_1\bar{x}_B\omega)\,S_p(\bar{x}_A\bar{x}_B\omega)\,d\bar{x}_A\,d\bar{x}_B \qquad (27)$$

APPROXIMATE VALUE OF MEAN SQUARE RESPONSE LEVEL

The mean square displacement $\overline{w^2(\bar{x}_1 t)} = \int_{-\infty}^{\infty} S_w(\bar{x}_1\omega)\,d\omega$. For lightly damped modes well separated in frequency we can neglect the cross terms. We can therefore make an estimate of the mean square overall level by performing the frequency integration for each mode separately. Use the result for a single degree of freedom case.

$$\overline{w^2(x_1,t)} \simeq \sum_{r=1}^{N} \frac{\pi}{\omega_r^3 \eta_r} \frac{\phi_r^2(\bar{x}_1)}{M_r^2} \int_A \int_A \phi_r(\bar{x}_A)\ \phi_r(\bar{x}_B)\ S_p(\bar{x}_A\ \bar{x}_B\ \omega_r)\ d\bar{x}_A\ d\bar{x}_B \qquad (28)$$

FURTHER CONSIDERATION OF THE RESPONSE EQUATION

The most general equation describing the response of structures to acoustic loads (equation (25)) can be investigated to see what approximations can be reasonably made in specific cases. Examples are also given of typical results for assumed pressure fields.

1. The cross terms

A considerable simplification can be achieved if the cross terms can be neglected. Let us now look at this possibility. Consider first the cross joint acceptance function and the extremes of spatial correlation.

Case 1 Fully correlated pressures i.e. plane wave normally incident on the structure.

$$S_p(\bar{x}_A\bar{x}_B\omega) = S_p(\omega)$$

Then

$$j_{rs}^2 = \frac{1}{A^2} \int_A \int_A \phi_r(\bar{x}_A)\ \phi_s(\bar{x}_B)\ d\bar{x}_A\ d\bar{x}_B$$

$$= \frac{1}{A^2} \int_A \phi\ (\bar{x}_A)\ d\bar{x}_A \int_A \phi_s(\bar{x}_B)\ d\bar{x}_B$$

In many structures the mode shapes will approach a sinusoidal shape as r increases. For sinusoidal shapes:

$$\int \phi_r(\bar{x}_A)\ dx_A = 0 \quad \text{for r even}$$
$$\to 0 \quad \text{as r increases for r odd.}$$

The first cross term $j_{r,r+1}$ will thus be zero. This will also be true of the third, fifth etc. The intermediate terms will in general be small but non zero and decrease as the separation increases.

Case 2 Spatially uncorrelated pressures

$$S_p(\bar{x}_A\bar{x}_B\omega) = S_p(\omega)\ \delta(\bar{x}_A - \bar{x}_B)$$

Then $$j_{rs}^2 = \frac{1}{A^2} \int_A \phi_r(\bar{x}_A)\ \phi_s(\bar{x}_A)\ d\bar{x}_A = 0 \qquad r \neq s$$ because of orthogonality.

Thus we see from these bounding results that the j_{rs}^2 cross term will be zero in all but a few cases when it may have a small value.

The second part of the cross term arises from the frequency function $(X_r - iY_r)(X_s + iY_s)$. The value of this term depends on the closeness of the natural frequencies and the size of the modal damping.

The net result of these two effects is that the contribution to the overall mean square level from the cross terms is generally zero or very small. Mercer (1) showed that in the case of a structure having many close natural frequencies the cross terms only contributed about 5% of the total rms level. In the worst case the contribution rose to about 20%. It is therefore proposed that we omit cross terms from the calculation.

2. Approximations for use in evaluating the direct terms

The response spectral density can now be written as

$$S_w(x_1\omega) = S_p(\omega)\ A^2 \sum_{r=1}^{\infty} \frac{\phi_r^2(x_1)}{M_r^2\ [(\omega_r^2 - \omega^2)^2 + \eta_r^2\omega_r^4]}\ j_{rr}^2$$

When the mode shape is sinusoidal or approaches a sinusoidal shape at high mode numbers $M_r \simeq M \times \frac{1}{2}$ if the mass is uniformly distributed [M is the total mass of the structure].

With the exception of j_{rr}^2, reasonable estimates of all the other parameters can be obtained:

a) $S_p(\omega)$ Use measured spectra, empirically developed design spectra given in several of the design codes, or theoretical results which are available for a few cases.

The major practical difficulty arises when the level of the spectrum varies from place to place over the structure. White(2) has discussed this aspect and suggests the use of the spatially averaged spectrum. This will probably be adequate except in extreme cases such as tall chimneys whose height approaches the earth's boundary layer thickness.

b) $\phi_r(x_1)$ This is the modal displacement at the point of interest x_1 .

c) M_r Can be approximated by $\frac{1}{2}M$ except at perhaps the first two or three modes or where there are large concentrated masses on the structure.

d) ω_r Computed natural frequencies of the structure. In vacuo values will usually be adequate.

e) η_r A value of about 2% is typical for many engineering structures. If better information is not available from measurements 2% can be used as a default value. The major exception is structures where there is significant fluid loading such as offshore structures. In such cases a higher value is more appropriate.

3. The joint acceptance function j_{rr}^2

The coupling coefficient or joint acceptance function is a measure of the efficiency with which a particular pressure field excites a specific mode of vibration of the structure. Analytical results can be obtained for three classes of general pressure field. In most practical cases one of these results can be used to give an estimate of the response. The three classes of pressure field are:

1. A 'frozen' convected field
2. A convected field with statistical decay
3. A reverberant pressure field

Case 1 has already been considered (eqn. 12). This represents the first approximation to such excitations as boundary layer turbulence, atmospheric turbulence, deep water sea motion etc. A better approximation to these conditions in convected turbulence is Case 2. In this approximation the cross spectral density function is represented in the direction of the convection by a decaying cosine wave:

$$S_p(x_A x_B \omega) = S_p(\omega)\, e^{-\alpha \frac{\omega}{c}(x_A - x_B)} \cos \frac{\omega}{c}(x_A - x_B)$$

where x_A and x_B are distances along the direction of convection and c is the velocity of convection.

Some typical experimental results are shown in figure 4 for boundary layer pressure fluctuations (3) and in figure 5 for near field jet noise pressures (4). Powell (5) has proposed a method of handling the computation for a two or three dimensional surface where the effective convection velocity or scale is different in the different directions across the surface.

To illustrate the effect of the parameters c and α computations have been made on the case of a simply supported beam excited by a pressure field convecting along its length. The effect of the decay parameter α on the joint acceptance function is illustrated in figure 6. These show that in the lowest mode (r=1) the decay has very little effect except at the higher frequencies. For the higher modes the peak value of the joint acceptance increases slightly and also the size of the side lobes increases as the decay parameter increases.

The reverberant sound field has a cross spectral density function of the form

$$S_p(x_A x_B \omega) = \frac{\sin \frac{\omega}{c}(x_A - x_B)}{\frac{\omega}{c}(x_A - x_B)}\, S_p(\omega)$$

The joint acceptance for this function (from ref. 6) is shown in figure 7. Curves chosen from one of these three types of pressure field can be used to give estimates of the joint acceptance function in many cases of practical interest. An upper bound for the joint acceptance has been derived by White (7) and shown to be close to the exact results obtained by Bozich (8) for a decaying cosine cross spectral density function.

SOME EXPERIMENTAL RESULTS

Many measurements have now been taken of displacement or strain response of structures which are excited by random pressures. However in few of these cases has there been any computation of the response for comparison with theory. This has generally happened because either all the input parameters are not known or else the computation presented too formidable a problem. Perhaps the most completely documented cases concern boundary layer excitation of aircraft structures. The comparisons have been made on laboratory tests in wind tunnels where the flow characteristics and the structural vibration characteristics are known or have been measured accurately. Examples of the measured and computed response spectra for a single fully fixed panel are shown in figure 8 (9). These comparisons show that good agreement can be achieved in the lower modes. The accuracy is less good in the higher modes. The response in these modes is much less than that in the fundamental and therefore the discrepancy is of less practical significance.

APPROXIMATIONS BASED ON A SINGLE MODE RESPONSE

In some cases (10, 11) it has been shown that a reasonable approximation can be obtained by making the assumption that the majority of the response is in the first mode of the structure. This seems to give a satisfactory estimate in the case of relatively large plate like structures subjected to broad band acoustic loads.

Assuming only one significant mode of response, the mean square level of displacement from equation (28) is:

$$\overline{w^2(t)} = \frac{\pi}{\omega_r^3 \eta_r} \frac{\phi_r^2(x_1)}{M_r^2} \int_A \int_A \phi_r(x_A)\, \phi_r(x_B)\, S_p(x_A x_B \omega_r)\, dA\, dA \qquad (29)$$

where r is the predominant mode - usually the first mode. This now eliminates the mode summation and simplifies the structure to a set of independent plates responding in their fundamental mode only.

The second major simplification is to assume that the pressures are exactly in phase over the whole plate. The equation then becomes:

$$\overline{w^2(t)} = \frac{\pi}{\omega_r^3 \eta_r} \frac{\phi_r^2(x_1)}{M_r^2} \left\{ \int_A \phi_r(x_A)\, dA \right\}^2 S_p(\omega_r) \qquad (30)$$

This can be written in terms of the displacement response of the plate to a uniform static pressure of unit magnitude. The static displacement w_o at x_1 is given by:

$$w_o = \frac{\int_A \phi_r(x_A)\, dA}{\omega_r^2 M_r} \cdot \phi_r(x_1)$$

Thus equation (30) can be written as:

$$\overline{w^2(t)} = \frac{\pi}{\eta_r} \omega_r S_p(\omega_r) w_o^2 \qquad (31)$$

The more usual expression for the mean square stress can be written in terms of the viscous damping ratio ξ and frequency in cycles per second f, as:

$$\overline{\sigma^2(t)} = \frac{\Pi}{2\xi} f_r S_p(f_r) \sigma_o^2 \qquad (32)$$

where σ_o is the stress at the point of interest due to a uniform static pressure of unit magnitude. This expression was first derived from the consideration of a single degree of freedom system by Miles (12).

THE TRAVELLING WAVE METHOD

In many physical situations the acoustic pressures acting on a structure are created by a wave which travels along the surface of the structure. Wave motion is then induced in the structure itself. The waves propagate in the same direction as the acoustic waves which generate them, but on reaching a boundary or discontinuity in the surface they are partly reflected, partly transmitted across the boundary, and also possibly partly transformed into different types of waves.

When the reflections interact with the incident structural waves, standing waves are established and at particular frequencies these constitute the normal modes of vibration of the surface. It is particularly helpful in problems of acoustically excited vibration to see the vibration from the wave rather than the mode viewpoint. A better understanding of the problem may often thus be obtained, and certain specific problems are more easily solved using wave concepts.

Plane flexural wave motion in uniform plates

A plane wave is one in which the wave displacement at any instant is identical at all points along a line perpendicular to the direction of propagation. In this section we consider such waves propagating parallel to the x axis of an infinite uniform plate.

When a flat plate is subjected to a transverse, time dependent pressure loading, p(x,y,t), the transverse deflection w is governed by the fourth order differential equation.

$$D \nabla^4 \bar{w} = p$$

where

$$\nabla^4 = \left\{ \frac{\partial^4}{\partial x^4} + 2 \frac{\partial^4}{\partial x^2 \partial y^2} + \frac{\partial^4}{\partial y^4} \right\} \qquad (33)$$

and

$$D = Eh^3/12(1 - \nu^2) .$$

The transverse loading under free wave conditions stems entirely from inertia loading, $-\rho h \partial^2 \bar{w}/\partial t^2$, so equation (33) becomes the flexural wave equation

$$D \nabla^4 \bar{w} + \rho h\, \partial^2 \bar{w}/\partial t^2 = 0 \qquad (34)$$

where ρ is the plate density. When the wave motion is harmonic we express $\bar{w}$ in the form

$$\bar{w} = w e^{i\omega t} \qquad (35)$$

in which w is a function of x and y only.

The differential equation for w, now becomes

$$\nabla^4 w - k^4 w = 0 \qquad (36)$$

where $k^4 = \rho h \omega^2 / D$

This is a general equation, applicable to all forms of the plate flexural wave motion, and not just to the plane waves parallel to the x axis. However, the latter wave motion, being independent of the y co-ordinate is represented by the solution of (36) in the form

$$w = A e^{k_n x}$$

where k_n can have any one of the four roots of

$$k_n = \{\rho h \omega^2 / D\}^{\frac{1}{4}}$$

Let the real positive root of this be denoted by k. The four different k_n's are $k_1 = k$, $k_2 = -k$, $k_3 = ik$, and $k_4 = -ik$. The complete solution of the wave equation includes all these k's in the form

$$\bar{w} = A_1 e^{kx} e^{i\omega t} + A_2 e^{-kx} e^{i\omega t} + A_3 e^{i(\omega t + kx)} + A_4 e^{i(\omega t - kx)} \qquad (37)$$

The last term of this expression represents a true flexural wave, sinusoidal in spatial wave-form, constant in amplitude (A_4) and propagating to the right (in the positive x direction). In this term k represents the difference in phase between deflections at points which are separated by unit x-wise distance apart, and is the flexural wave number. The wavelength of the flexural wave is given by

$$\lambda = 2\pi/k = 2\pi \{D/\rho h \omega^2\}^{\frac{1}{4}} . \qquad (38)$$

The flexural wave velocity, a_f, is given by λ x frequency (Hz) so

$$a_f = \omega^{\frac{1}{2}} \{D/\rho h\}^{\frac{1}{4}} . \qquad (39)$$

Now the longitudinal wave velocity in the flat plate, a_ℓ, is given by $\{E/\rho(1-\nu^2)\}^{\frac{1}{2}}$. Also, $h^2/12 = \kappa^2$, where κ = radius of gyration of the plate section. Hence

$$a_f = \{\omega a_\ell \kappa\}^{\frac{1}{2}} . \qquad (40)$$

This increases with (frequency)$^{\frac{1}{2}}$, so the waves of different frequency travel at different speeds. The wave motion is therefore dispersive.

The third term of equation (37) represents a similar wave which propagates to the left. Each of these waves (A_3 and A_4) transmits energy along the beam (in opposite directions in each case). The energy flow across a given plate section per unit time is found from the average value of (Shear force, S_x, on the section due to the wave x transverse plate velocity) + (Bending moment, M_x, on the section x angular plate velocity).

These two component energies are found to be equal in a free wave, and the total power transmitted by a single propagating flexural wave of amplitude $|A|$ is

$$\Pi_f = Dk^3 \omega |A^2| = \omega^2 \rho h a_f |A^2| \qquad (41)$$

per unit width of plate, normal to the wave direction.

The first two terms of equation (37) do not represent propagating waves. The motion represented by $A_1 e^{kx} e^{i\omega t}$ is in phase at all points along the beam, and is a deflection which decays exponentially in space from right to left. $A_2 e^{-kx}$ represents a similar deflection which decays from left to right. No energy is transmitted by these 'waves', so we refer to them as 'non-propagating' waves. They always decay with distance away from the point at which they are generated, and will not be noticed (due to the exponential factor) at large distances from the point. The propagating waves, on the other hand, have the same amplitude at large or small distances, provided there is no damping present to attenuate them. Since the non-propagating waves are only noticed near to the point of generation, they may be called 'near-field' waves, whereas the propagating waves spreading outwards without attenuation, may be called 'far-field' waves.

The forced vibrations of a uniform flat plate

Consider a plate which is simply-supported along the two x wise edges, and which is subjected to a propagating pressure wave on the surface of the form

$$p(x,y,t) = p_o e^{i(\omega t - k_p x)} \qquad (42)$$

where k_p is the wave-number of the pressure wave on the plate surface. This represents a plane harmonic pressure wave propagating in the x direction only at a velocity of

$$a_p = \omega/k_p \quad .$$

At a given instant and x position, p is constant in the y direction. It is convenient, however, to analyse this constant value into sinusoidal y wise components, thus:

$$P_o = \sum_{r=1}^{\infty} P_r \sin(2r-1)\,\pi y/\ell_y = \sum_{r=1}^{\infty} P_r \sin(k_{ry} y) \quad . \tag{43}$$

For simplicity in the remainder of this analysis, we shall consider only the first term of this, i.e.

$$p(x,y,t) = P_1 \sin(k_{1y} y)\, e^{i(\omega t - k_p x)} \quad . \tag{44}$$

Now the response of the plate, $\bar{w}(x,y,t)$, to this pressure is governed by

$$D\nabla^4 \bar{w} - \omega^2 \rho h \bar{w} = P_1 \sin(k_{1y} y)\, e^{i(\omega t - k_p x)} \tag{45}$$

Damping may now be allowed in the plate by assinging the complex form $D(1+i\eta)$ to the flexural rigidity. The solution to equation (45) has two parts - the particular integral and the complementary function. The particular integral is identical to the response of the plate if it was infinite in the x direction, and represents a forced flexural wave travelling in the x direction. It is given by

$$w_{PI} = P_1 \sin(k_{1y} y)\, e^{i(\omega t - k_p)} \;/\; (D(k_{1y}^2 + k_p^2)^2 - \omega^2 \rho h) \quad . \tag{46}$$

Since the plate is in fact finite, this wave will be reflected when it encounters the right hand boundary of the plate. Likewise, reflection will occur from the left hand boundary. The reflections, however, will be free flexural waves and there will be a propagating wave and a near field wave reflected from each boundary. These free waves constitute the complementary function of the solution to equation (45) and are represented by

$$w_{CF} = \{ A_1 e^{k_n x} + A_2 e^{-k_n x} + A_3 e^{ik_x x} + A_4 e^{-ik_x x} \}\, e^{i\omega t} \sin(k_{1y} y) \tag{47}$$

where

$$k_n = \{ k^2 + k_{1y}^2 \} \;,\; \{k_x = k^2 - k_{1y}^2 \} \;,\quad k^4 = \omega^2 \rho h / D \quad .$$

The total displacement at any point (x,y) is therefore given by

$$w(x,y,t) = \{A_1 e^{k_n x} + A_2 e^{-k_n x} + A_3 e^{+ik_x x} + A_4 e^{-ik_x x} + P_1 e^{-ik_p x} \;/\; \{D(k_{1y}^2 + k_p^2) - \omega^2 \rho h\}\}\sin(k_{1y} y) e^{i\omega t} \quad . \tag{48}$$

The solution containes four unknown constants of integration which can be found by ensuring that the solution satisfies the four boundary conditions of the plate at $x = 0$ and $x = \ell_x$. Suppose these two boundaries are encastré i.e. $w = 0$, $w' = 0$, at $x = 0$ and $x = \ell_x$.

These yield the following four equations for the A's.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ k_n & -k_n & ik_x & -ik_x \\ e^{k_n \ell_x} & e^{-k_n \ell_x} & e^{ik_n \ell_x} & e^{-ik_x \ell_x} \\ k_n e^{k_n \ell_x} & -k_n e^{-k_n \ell_x} & ik_x e^{ik_x \ell_x} & -ik_x e^{-ik_x \ell_x} \end{bmatrix} \begin{Bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{Bmatrix} = \begin{Bmatrix} 1 \\ -ik_p \\ e^{-ik_p \ell_x} \\ -ik_p e^{-ik_p \ell_x} \end{Bmatrix} \times \frac{P_1}{D(k_{1y}^2 + k_p^2) - \omega^2 \rho h} \tag{49}$$

Notice that the A's will have very large values under two different conditions:

(a) When the real part of the denominator of the right hand side approaches zero, i.e. when $\mathrm{Re}\,(D(k_{1y}^2 + k_p^2) - \omega^2 \rho h) \to 0$. To a first approximation this means $k_p \to \sqrt{k^2 - k_{1y}^2} = \mathrm{Re}(k_x)$.

Hence, when the wave number of the exciting pressure field is equal to the natural wave number of free flexural wave motion, a large response is generated. These two wave numbers will be equal if the propagation velocity of the pressure field is equal to the free wave velocity of the corresponding flexural waves in the plate at that frequency.

(b) When the determinant of the matrix on the left hand side has a minimum value. This occurs whenever the frequency is equal to one of the natural frequencies of the finite plate.

The response expression (48) will therefore give rise to resonant type peaks at frequencies corresponding to each of these conditions. Condition (a) is the 'COINCIDENCE CONDITION', for the propagation speed of the pressure wave coincides with the natural propagation velocity of the flexural wave being generated. Condition (b) is a simple 'resonance' condition.

Now a set of equations of identical form to (49) exists for each value of k_{ry} belonging to the pressure field. The equations can be solved for each of these values and for given values of p_o, k_p and ω. Hence, the displacement at any point on the plate can be found for a given value of k_p and ω. We can therefore write

$$w(x,y,t) = Y(x,y,k_p,\omega) \; P_o e^{i\omega t} \qquad (50)$$

where $Y(x,y,k_p,\omega)$ is the 'wave receptance function'. It is the complex harmonic plate response at point (x,y) to a harmonic pressure wave of given y-wise form of unit amplitude and given frequency and wave number.

The random vibration of uniform flat plates

The normal method (as outlined in the first part of this paper) can be applied to any type of plate, whether flat or curved, rectangular or irregular in shape, uniform or variable in thickness. The general expression obtained for the response spectral density applies to all such conditions. It involves the calculation of the joint acceptances and this can be an extremely tedious task; so also can be the summing of the modal responses allowing for the correlation between different modes.

Some simple types of structure can be more easily analysed by making use of a wave approach developed from the method described above. The uniform rectangular flat plate is one such case.

A random pressure field can be regarded as a continuous assembly of harmonic pressure waves of all different frequencies and wavelengths (or wave-numbers). A random pressure field which is convected along at a uniform velocity a_p and which maintains its precise wave-form as it moves along can be analysed into its continuum of spectral components of all frequencies. Each frequency can be associated with a unique wavelength of pressure wave (given by $\lambda = a_p/f = 2\pi a_p/\omega$) or with a unique wave number (given by $k_p = \omega/a_p$). Thus, the spectral density $S_p(\omega)$ of the pressure fluctuation corresponds to a wave number $k_p = \omega/a_p$.

If the random pressure field does not maintain its precise waveform as it convects along (and this is the more realistic case) it can be analysed into a continuous assembly of harmonic pressure waves of all frequencies and all wave numbers at each frequency. Physically, this means that we have pressure waves of all different frequencies travelling along at all different velocities. The spectral information about such pressure waves can be described by the two-dimensional wave number /frequency spectrum. For any given frequency, a spectrum is drawn in the wave number domain showing how the pressure 'power' is distributed over the different wave numbers.

Now the power spectrum of the pressure at a single point can be found from the Fourier transform of the pressure auto-correlation function, i.e.

$$S_p(\omega) = (1/2\pi) \int_{-\infty}^{+\infty} R_p(\pi) \; e^{-i\omega t} \, d\tau$$

where $R_p(\tau)$ is the pressure auto-correlation function with time-delay τ and is defined by

$$R_p(\tau) = \lim_{T\to\infty} (1/2T) \int_{-T}^{+T} p(t) \; p(t+\tau) \, dt \; .$$

We can also define the space-time correlation function as the correlation between the pressure, p(x,t) at one point and the pressure $p(x+\xi, t+\tau)$ at another point distance ξ (in the x direction) from the first point. This is

$$R_p(\xi,\tau) = \lim_{T\to\infty} (1/2T) \int_{-T}^{+T} p(x,t) \; p(x+\xi,t+\tau) \, dt$$

The wave number/frequency spectrum $S_p(k_p,\omega)$ is then defined as the double Fourier Transform of the space-time correlation function, i.e.

$$S_p(k_p,\omega) = 1/(2\pi)^2 \int_{-\infty}^{+\infty} e^{-ik_p\xi} \int_{-\infty}^{+\infty} R_p(\xi,\tau) \; e^{-i\omega\tau} \, d\tau \, d\xi$$

Now the cross power spectral density of the pressures at any two points distance ξ apart is

$$S_p(\xi,\omega) = (1/2\pi) \int_{-\infty}^{+\infty} R_p(\xi,\tau) \; e^{-i\omega\tau} \, d\tau$$

This can be expressed in the form

$$S_p(\xi,\omega) = S_p(\omega)\ \rho(\xi,\omega)$$

where $\rho(\xi,\omega)$ is the narrow band correlation coefficient between pressures distance ξ apart.

Hence,

$$S_p(k_p,\omega) = S_p(\omega)\ \frac{1}{2\pi} \int_{-\infty}^{+\infty} \rho(\xi,\omega)\ e^{-ik_p\xi}\ d\xi$$

The total power spectral density of the pressure at frequency ω is found from this by integrating over the whole range of k_p. Thus

$$S_p(\omega) = \int_{-\infty}^{+\infty} S_p(k_p,\omega)\ dk_p$$

and the mean square value of the pressure at a point is found by integrating this over the whole of the frequency range, hence

$$\langle p^2 \rangle = \int_{-\infty}^{+\infty} S_p(\omega)\ d\omega.$$

Now the spectral density of the plate displacement is related to the spectral density of the pressure by

$$S_w(k_p,\omega) = |Y(x,y,k_p\omega)|^2\ S_p(k_p,\omega)$$

where $Y(x,y,k_p,\omega)$ is the wave receptance function of equation (50). The frequency spectrum of the displacement ^{p} is obtained by integrating this over the whole wave number domain, so

$$S_w(\omega) = \int_{-\infty}^{+\infty} |Y(x,y,k_p,\omega)|^2\ S_p(k_p,\omega)\ dk_p$$

and the mean square displacement is

$$\langle w^2 \rangle = \int_{-\infty}^{+\infty} S_w(\omega)\ d\omega.$$

Further reading on the Travelling Wave Method

L. Cremer, M. Heckl, E.E. Ungar, 1973, Structure-borne Sound, Springer-Verlag, New York, Heidelburg, Berlin.

E. Skudrzyk, 1968, Simple and Complex Vibratory Systems, Pennsylvania State University Press, University Park and London.

REFERENCES

1. C.A. Mercer, 1965, J. Sound Vib. Vol 2 p.293. 'The response of a multi supported beam to a random pressure field'.

2. P.H. White, ASME Paper 70 WA/DE-11, 'Response of Structures to Non Homogeneous Random Pressure Fields'.

3. M. Harrison, 1958, David Taylor Model Basin Report 1260, 'Pressure fluctuations on the wall adjacent to a turbulent boundary layer'.

4. B.L. Clarkson, 1965, Acoustical Fatigue in Aerospace Structures, Syracuse University Press, 'Scaling of the near field pressure correlation patterns around a jet exhaust'.

5. A. Powell, 1964, J. Acoust. Soc. Am. Vol 36 p783, 'On the estimation of the generalised force due to random pressure and on necessary modes'.

6. R.L. Barnowski, J.R. Maurer, 1970, NASA CR 1660, 'Distributed System Response Characteristics in Random Pressure Fields'.

7. P.H. White, 1964, J. Acoust. Soc. Am. Vol 36 pp 784-785, 'Some useful approximations for determining the vibration of structures excited by random pressures'.

8. D.J. Bozich, 1964, J. Acoust. Soc. Am. Vol 36 No.1 pp 52-58, 'Spatial Correlations in Acousto-Structural Coupling'.

9. D.R. Blackman, D.M. Clark, G.J. McNulty, J.F. Wilby, AFFDL TR 67-97, 'Boundary Layer Pressure Fluctuations and Structural Response'.

10. B.L. Clarkson, 1968, Aero Journal, 72, pp 1000-1010, 'Stresses in skin panels subjected to random acoustic loading'.

11. A.G.R. Thomson, Acoustic Fatigue Design Data, Parts 1,2,3 & 4, AGARDograph No. 162.

12. J.W. Miles, 1954, J. Aero. Sci. Vol 21 pp 753-762, 'On Structural Fatigue under Random Loading'.

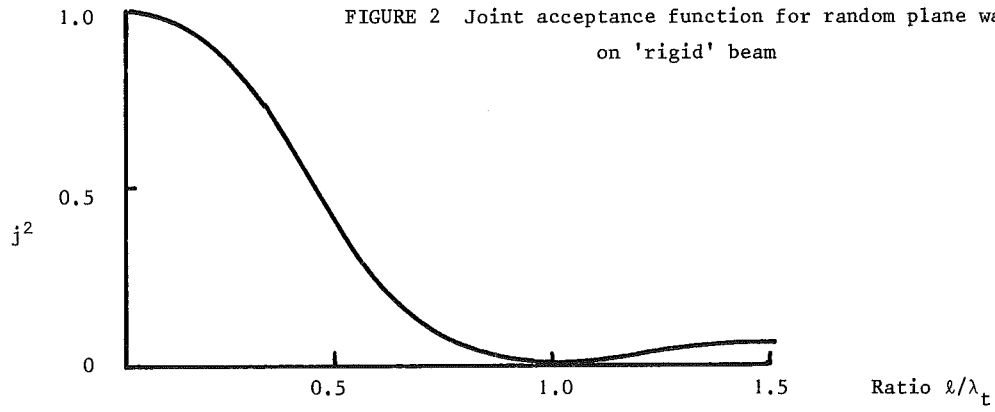


FIGURE 2 Joint acceptance function for random plane waves on 'rigid' beam

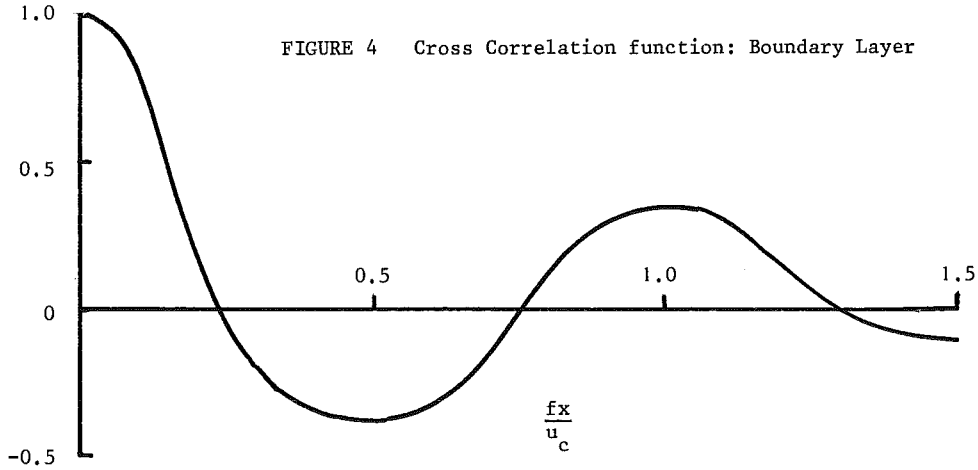


FIGURE 4 Cross Correlation function: Boundary Layer

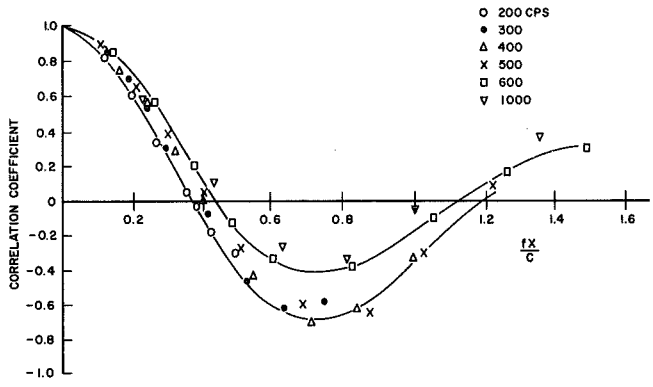


FIGURE 5 Cross correlation function for near field jet noise pressures

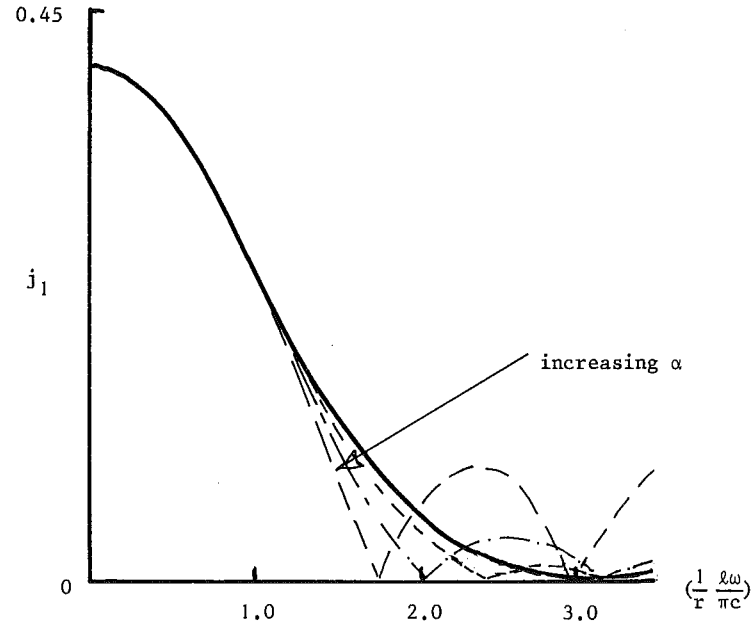


FIGURE 6 Joint Acceptance Function for:

$$S_p(x_A x_B \omega) = S_p(\omega)\, e^{-\alpha \frac{\omega}{c}(x_A - x_B)} \cos \frac{\omega}{c}(x_A - x_B)$$

Simply supported mode shape $\phi(x) = \sin \frac{r\pi x}{\ell}$

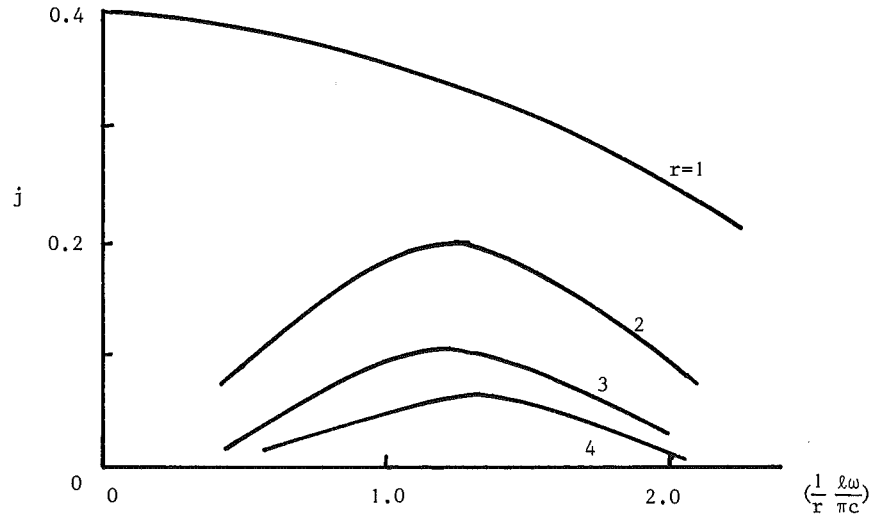


FIGURE 7 Joint Acceptance Function for a reverberant field

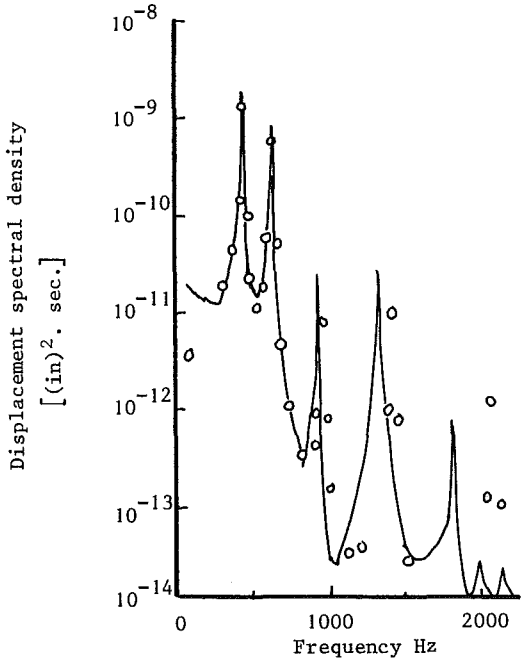


FIGURE 8 Deflection Spectral Density of panel excited by Boundary Layer Pressure (Maestrello)

FUNDAMENTAL CONCEPTS OF SOUND RADIATION

M. C. Junger
Cambridge Acoustical Associates, Inc.
54 Rindge Avenue Extension
Cambridge, Massachusetts 02140, USA

It is the function of this lecture to recall the fundamental physical concepts underlying sound radiation. The wave equation is derived and its solution constructed for plane waves. Impedance and wave number concepts are introduced. The sound field for a uniformly pulsating spherical radiator and its limiting case, the point source, is derived. Finally, the scaling laws of acoustical model experiments are stated.

Sound is a space- and time-dependent deviation of pressure from the static pressure. These pressure fluctuations come about as a result of vibrations of a boundary in contact with the acoustic fluid, of time-dependent forces or stresses applied to this fluid, of pulsating fluid injection, and of oscillatory or transient thermal expansion of the fluid.

The concepts presented here are applicable to sound no matter how it is originated. They are therefore relevant to flow-generated noise as well as to sound radiated by harmonically vibrating structures or transducers, within the constraints of the assumptions inherent in our simple mathematical model.

A. THE LINEARIZED WAVE EQUATION AND ITS PHYSICAL INTERPRETATION

1. Mathematical Model of the Acoustic Fluid

Viscous stresses in the acoustic fluid are assumed negligible compared to inertial and elastic stresses. Any convection velocity is similarly assumed small compared to the sound velocity. The principal assumption, which permits linearization of the equations, is that sound pressure, i.e., the deviation $dP_s \equiv P$ from mean pressure P_s, is small compared to this mean pressure. The validity of this assumption for the majority of practical situations is readily verified: For air-borne sound, the threshold of pain is 120 dB re 0.0002 μbar (or, in MKS units, 144 dB re 1 μPa and, in the CGS system which we shall mostly use, 44 dB re 1 μbar). Atmospheric pressure is 20 log 10^6 = 120 dB re 1 μbar. The difference is therefore four orders of magnitude.

For an exhaustive analysis of various effects bearing on sound propagation, the reader is referred to Hunt's derivation of the wave equation.[1]

2. The Equation of Motion

The localized spatial distribution of sound-associated pressure distributions results in acceleration of fluid particles towards regions of lesser pressure. From Newton's Second Law,

$$\nabla p = -\rho \partial^2 \bar{d}/\partial t^2 \tag{1}$$

where $\bar{d}$ is a vector displacement and ρ the fluid density. This relation is valid whether the fluid is incompressible or not.

3. The Bulk Modulus

The 3-dimensional equivalent of Hooke's Law relates the fractional density change to the pressure change:

$$B \equiv p/(d\rho/\rho) \tag{2}$$

It is convenient to eliminate density in terms of volume strain. From conservation of mass considerations,

$$d(V\rho) = Vd\rho + \rho dV = 0$$

or

$$d\rho/\rho = -dV/V = -\nabla \cdot \bar{d}$$

Eq. 2 now becomes

$$B = -p/(dV/V) \tag{3}$$

The linearized volume strain is conveniently expressed in terms of the divergence of the fluid particle displacement:

$$p = -BV \cdot \bar{d} \tag{4}$$

Except for unusually large ratios of frequency to pressure, sound waves compress gases nearly adiabatically:

$$d(P_s V^\gamma) = pV^\gamma + P_s V^{\gamma-1} dV = 0$$

where $\gamma = 1.4$ for air. Solving for volume strain and substituting the result in Eq. 3,

$$B = \gamma P_s \tag{5}$$

For air under representative atmospheric conditions, $B = 1.42 \times 10^6$ μbar. For water, $B = 2.13 \times 10^{10}$ μbar.

4. The Wave Equation

One of the two unknown variables, for example the fluid particle displacement vector, can be eliminated between Eqs. 1 and 4. For this purpose we take the divergence of Eq. 1 and invert the order of differentiation,

$$\nabla^2 p = -\rho \frac{\partial^2}{\partial t^2} (\nabla \cdot \bar{d})$$

Furthermore, we differentiate Eq. 4 twice with respect to time:

$$\frac{\partial^2 p}{\partial t^2} = -B \frac{\partial^2}{\partial t^2} (\nabla \cdot \bar{d})$$

Eliminating $\partial^2(\nabla \cdot \bar{d})/\partial t^2$ between these two equations, we obtain the linearized wave equation:

$$\nabla^2 p - \frac{\rho}{B} \frac{\partial^2 p}{\partial t^2} = 0 \tag{6}$$

Alternatively, we could have constructed the wave equation in terms of fluid particle displacement, or in fact fluid density. The wave equation was first derived by d'Alembert to describe the vibrations of a string (1747). The wave equation establishes the relation between fluctuations in space and in time. If the medium is assumed incompressible ($B = \infty$), the temporal adjustment takes place instantaneously. Eq. 6 degenerates to the Laplace equation:

$$\nabla^2 p = 0 \tag{7}$$

B. PLANE SOUND WAVES

1. The Sound Velocity

Let us seek a one-dimensional solution to Eq. 6, whereby the pressure is dependent only on one spatial coordinate and time. Sound waves are therefore plane, such as would be observed at large range from their source. For this situation, the wave equation can be approximated as

$$\frac{\partial^2 p}{\partial x^2} - \frac{\rho}{B} \frac{\partial^2 p}{\partial t^2} = 0 \tag{8}$$

This equation is satisfied by functions of the form

$$p(x,t) = F\left[x \mp \left(\frac{B}{\rho}\right)^{1/2} t\right] \tag{9}$$

The coefficient multiplying time has units of velocity. An observer travelling with speed

$$c = (B/\rho)^{1/2} \tag{10}$$

sees no change in sound pressure. This is therefore the sound velocity. For a gas, Eq. 5 holds:

$$c = (\gamma P_s/\rho)^{1/2} \tag{11}$$

This result was derived by Laplace (1816). Since t always increases, the negative sign in Eq. 9 corresponds to sound propagating in the positive x-direction, and vice versa. Previously, Newton (1687) had constructed the isothermal sound velocity obtained by setting $\gamma = 1$ in Eq. 11. For representative conditions, the sound velocity in air is 344 m/s, and in water 1,460 m/s.

2. The Helmholtz Equation

This presentation will now be specialized to a harmonic time dependence. The expression describing one-dimensional pressure waves now takes the form

$$p(x,t) = \mathrm{Re}[p(x)e^{-i\omega t}] \tag{12}$$

For the sake of brevity, both the symbol Re and sometimes the harmonic function of time itself will be suppressed. Time-harmonic solutions are important not only in that they represent pure tones but also as building blocks of the statistical formulation of random sound pressures. Noting that for harmonic time dependence, the time-differential operator in Eq. 8 is tantamount to multiplication by $-\omega^2$, and combining Eqs. 8, and 10, the one-dimensional wave equation thus specialized becomes an ordinary differential equation:

$$\frac{d^2p}{dx^2} + \frac{\omega^2}{c^2} p = 0 \tag{13}$$

This is the Helmholtz equation. For the sake of brevity, the coefficient multiplying the pressure is defined in terms of the wavenumber

$$k \equiv \omega/c \tag{14}$$

3. One-Dimensional Pressure Fields

Solutions of Eq. 13 are in the form $P\exp(\pm ikx)$ or of combinations thereof, i.e., of trigonometric functions. The expression in Eq. 12 now becomes

$$p(x,t) = P\exp(\pm ikx - i\omega t) \tag{15}$$

where it is understood that only the real component $\cos(\pm kx-\omega t)$ is physically meaningful. The positive sign implies propagation in the positive x-direction and vice versa. At a given time t, the same pressure is encountered at locations $x+(2\pi n/k)$, where n is an integer. The interval corresponding to n=1 defines the wavelength $\lambda = 2\pi/k$.

4. Boundary Conditions

So far attention has been focussed on sound propagation in free space. Now consider the sound field near a boundary surface S described by the distribution of the normal unit vector $\bar{n}$ taken positive into the fluid medium. If we ignore viscosity, as we have throughout, there need not be continuity between tangential motions of the boundary and of fluid particles. There must however be continuity between the normal displacement w or acceleration $\ddot{w}$ of the boundary surface S and of the fluid particles. Taking the $\bar{n}$-component ofthe vector equation in Eq. 1, this boundary condition becomes

$$\frac{\partial p}{\partial n} = -\rho\ddot{w} = \rho\omega^2 w \text{ on } S \tag{16}$$

For a radiating surface endowed with a prescribed dynamic configuration w, this constitutes a non-homogeneous boundary condition which determines the amplitude of radiated pressure. This will be illustrated in Section C. For a rigid scatterer, $w \equiv 0$ and consequently,

$$\partial p/\partial n = 0 \quad \text{on } S \tag{17}$$

5. The Characteristic Impedance

For this simplest of dynamic configurations, the plane-wave solution in Eq. 15, vector notation is not required to describe the fluid particle displacement, which can therefore be denoted by the scalar w. The particle acceleration and velocity obtained by substituting Eq. 15 in Eq. 16, are

$$\frac{\partial p}{\partial x} = ikp = -\rho\ddot{w} = i\omega\rho\dot{w}$$

The characteristic impedance is defined as the ratio of pressure to particle velocity in a plane wave. From the above equation,

$$z \equiv \frac{p}{\dot{w}} = \frac{i\omega\rho}{ik} = \rho c = (B\rho)^{1/2} \tag{18}$$

where use has been made of the definition in Eq. 14. This same result can be derived for non-harmonic situations described by Eq. 9. The fluid particle acceleration is

$$\ddot{w} = -\frac{1}{\rho}\frac{\partial F}{\partial x} = \frac{1}{\rho}\left(\frac{\rho}{B}\right)^{1/2}\frac{\partial F}{\partial t}$$

Integrating with respect to time, and setting F = p, one obtains the result in Eq. 18. The characteristic impedance equals 42 μbar/cm/s for air and 1.5×10^5 in the same CGS units for water. Consequently, similar boundary motions produce pressures between three and four orders of magnitude larger in water than in air.

C. THE SPHERICAL SOURCE AND ITS IMPEDANCE

1. The Helmholtz Equation in Spherical Coordinates

The wave equation, Eq. 6, no longer restricted to plane waves but still specialized to harmonic situations is

$$\nabla^2 p + k^2 p = 0 \tag{19}$$

Now consider a source in the form of a uniformly pulsating sphere. The one-dimensional Laplace operator in the spherical radial coordinate system is

$$\nabla^2 \equiv \frac{d^2}{dR^2} + \frac{2}{R}\frac{d}{dR} \tag{20}$$

It can be verified that Eqs. 19 and 20 admit solutions of the form

$$p(R) = (Ae^{ikR}+Be^{-ikR})/R$$

For the time dependence exp(-iωt), the positive exponential indicates propagation in the direction of increasing R. Selecting these outgoing rather than converging waves, the pressure can be written explicitly as

$$p(R,t) = \frac{A}{R} e^{i(kR-\omega t)} \tag{21}$$

This differs from the plane-wave solution in that the pressure amplitude decreases with increasing range.

2. The Boundary Condition Specialized to Spherical Radiators

The undetermined coefficient in Eq. 21 can be expressed in terms of the velocity $\dot{W}\exp(-i\omega t)$ prescribed over the surface of the sphere. The boundary condition, Eq. 16 becomes

$$\frac{dp}{dR} = i\omega\rho\dot{W} \quad , \quad R = a \tag{22}$$

where 2a is the source diameter. From Eq. 21,

$$\frac{dp}{dR} = (ik - \frac{1}{a}) \frac{A}{a} e^{ika} \quad , \quad R = a \tag{23}$$

$$= (ika-1)p(a)/a$$

Combining Eqs. 22 and 23, one can solve for A and express the sound field in terms of the source velocity:

$$p(R) = \frac{\rho c(ka-i)ka^2\dot{W}}{(K^2a^2+1)R} \exp[ik(R-a)] \tag{24}$$

3. The Point Source

The point source (or simple source, or acoustic monopole) is defined as a uniformly pulsating sphere small in terms of wavelengths. Its pressure field, which will be used to construct more complicated sound fields, is identified here by the subscript s. The pressure is obtained by neglecting terms of order k^2a^2 compared to unity in Eq. 24:

$$p_s(R) = -\frac{ika^2\rho c\dot{W}}{R} \quad , \quad k^2a^2 << 1 \tag{25}$$

The surface pressure is in phase with the acceleration $\ddot{W} = -i\omega\dot{W} = -ikc\dot{W}$:

$$p_s(R) = \frac{\rho a^2\ddot{W}}{R} e^{ikR}$$

Noting that the volume acceleration is

$$\ddot{Q} = 4\pi a^2\ddot{W}$$

the pressure field of the point source can be written as

$$p_s(R) = \frac{\rho\ddot{Q}}{4\pi R} e^{ikR} \tag{26}$$

4. Impedance of Spherical Waves

The specific acoustic impedance of spherical waves is defined as

$$z(R) \equiv \frac{p}{\dot{w}} \tag{27}$$

Combing Eqs. 21, 22, and 23, this impedance becomes

$$z(R) = \frac{i\rho\omega\rho}{(dp/dR)} = \frac{i\rho ckR}{ikR-1} \tag{28}$$

It is convenient to express this complex impedance in terms of a reactance x and a resistance r

$$z = ix+r \tag{29}$$

It can be verified that

$$r = \frac{\rho c k^2 R^2}{k^2 R^2 + 1}$$

$$\simeq 0 \quad , \quad k^2R^2 << 1$$

$$\simeq \rho c \quad , \quad k^2R^2 >> 1$$

$$x = \frac{-\rho c k R}{k^2 R^2 + 1} \tag{30}$$

$$\simeq -\rho c k R \quad , \quad k^2R^2 << 1$$

$$\simeq 0 \quad , \quad k^2R^2 >> 1$$

Setting R=a, one obtains the impedance of a spherical source. This impedance is seen to be reactive at low frequencies and resistive at high frequencies.

5. Far-Field Criteria

In the above statement, "high frequencies" can be replaced by "ranges large in terms of wavelengths". Under these conditions, i.e., when $k^2R^2 >> 1$, the impedance equals the characteristic plane-wave impedance. This is one of the characteristics of the far-field. The other two are that the pressure decay with increasing range as R^{-1} and that the distribution-in-angle of the pressure be independent of range. The latter two conditions are satisfied at all ranges by the sound field of a uniformly pulsating sphere. This isthe only source configuration to which this applies, as will be seen in Lecture 7.

D. MODEL SCALING LAWS

The solution in Eq. 24 allows us to construct the laws governing scaled acoustic measurements, whereby a source of characteristic dimension L and amplitude of vibration W, is replaced by a model source of dimension L/s, of similar dynamic configuration but of reduced amplitude of vibration W/s, the acoustic fluid remaining unaltered. The desired scaling law insures that the pressure p_m measured in the model experiment at range R/s equals the full-scale pressure p at range R:

$$p_m\left(\frac{R}{s}\right) = p(R) \tag{31}$$

It is apparent that this relation is satisfied if ka and $\dot{W}$ remain invariant under scaling. Consequently, the model experiment must be conducted at frequency s . The scaling law therefore requires that linear dimensions including amplitudes of vibration normalized to wavelength be invariant. Since pressures and velocities both remain unchanged, so does the specific acoustic impedance.

REFERENCE

1 F. V. Hunt, "Propagation of Sound in Fluids", in D. E. Gray, Ed., American Institute of Physics Handbook, (New York, McGraw, Hill, 1957) pp. 325 et seq.

FUNDAMENTAL CONCEPTS OF FLOW-GENERATED NOISE

William K. Blake
Research Scientist
DAVID TAYLOR NAVAL SHIP
R&D CENTER
Bethesda, Maryland 20084
UNITED STATES

Summary

A summary of the fundamental properties of flow-generated sound and vibration will be given in this lecture. The elementary source types, the physical parameters governing sound and vibration, and general similarity rules will all be derived. The rudimentary effects of surfaces on flow-generated noise will also be reviewed.

1.0 Introduction

Flow-generated noise is a fully interdisciplinary subject area. The primary source of sound is some type of unsteadiness in the fluid generally resulting from unsteady momentum exchanges. Such fluid dynamics will impart time-varying forces to the surrounding fluid when the flow occurs around bodies and these forces generate unsteady reactive motions of the fluid as well as of the body at the boundary. When boundary motion occurs, one is often as much interested in the vibration of the body as in the sound produced. Indeed, sound resulting from the flow-induced body motion may be more important than the sound emitted from the forces acting on the fluid. Perhaps the most well known sound source involving flow past bodies is the Aeolian tone. Here a vortex street is formed behind a bluff body as illustrated schematically in Figure 1.

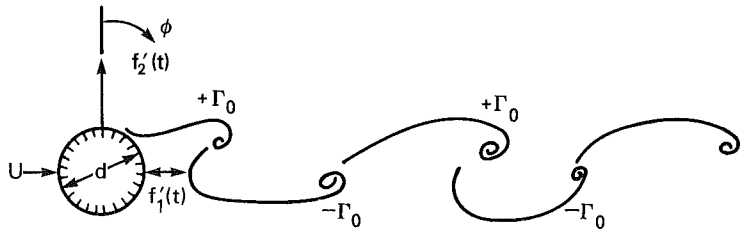


Figure 1 - Illustration of Vortex Street Behind Bluff Body and of Orientations of Unsteady Lift ($f_2'(t)$) and Drag ($f_1'(t)$) Forces. $f_1'(t)$ and $f_2'(t)$ Represent Forces Per Unit Axial Length.

Time-varying momentum exchanges accompanying the formation of the vortices of alternate circulation $+\Gamma_0$ cause periodic forces of equal and opposite sign to occur between the body and the fluid. Sound is emitted since those forces move both the fluid particles surrounding the body and possibly the body as well. The most well known noise source that does not involve a solid body vibration, but only unsteady momentum distributions in the fluid is jet noise. Here sound is created as a result of incomplete cancellations of fluid particle motions that accompany the interactions of turbulent eddies. In this lecture we will examine the fundamentals of sound generation by unsteady flows, and the concepts of dynamical of similitude governing each type. We will later consider in some detail the flow-generated sound produced by flow past bluff bodies and cylinders. Using the results of this analysis as a basis we will then present summaries of the governing parameters of some important practical noise sources.

2.0 The Cannonical Sources of Flow-Induced Sound

All sources of flow-induced sound are represented as some combination of three basic types which we will call cannonical. The most elementary type is the monopole source. It results from a net unsteady mass injection into the fluid region. Such a mass injection affects the surrounding fluid in all directions, it is therefore onmidirectional. Letting the rate of mass injection per unit volume be $\dot{q}(\vec{y},\omega)$ then the radiated sound pressure at a distance $r = |\vec{x}-\vec{y}|$ from the source region as illustrated in Figure 2, in a fluid medium in which reflecting surfaces are absent is

$$p_r(\vec{x},\omega) = \frac{1}{4\pi}\int_{\substack{\text{SOURCE}\\ \text{VOLUME}}} \dot{q}(\vec{y},\omega)\,\frac{e^{ik_0 r}}{r}\,dV(\vec{y}) \qquad (1)$$

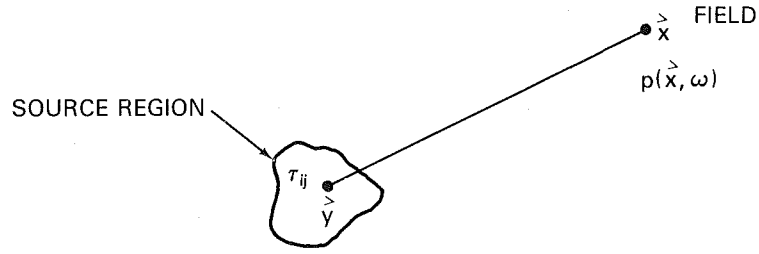


Figure 2 - Illustration of a Source Region and Field Point in an Unbounded Acoustic Medium.

where $k_o = \omega/c_o$. We shall assume throughout these lectures that all time-varying physical quantities may be represented by their Fourier transforms, e.g., for the acoustic pressure

$$p(\vec{x},\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\, p(\vec{x},t)\, dt \tag{2}$$

and

$$p(\vec{x},t) = \int_{-\infty}^{\infty} e^{-i\omega t}\, p(\vec{x},\omega)\, d\omega \tag{3}$$

This poses no practical limitation or loss of generality in the analysis. The formal implication of course is that the processes are steady state, so initial conditions are important, and temporally stationary in the statistical sense. Formal justification for the application of these integral formulae to random processes may be found in Cramer and Ledbedder[1]. In the case of random fluid processes $p(\vec{x},t)$ and $p(\vec{x},\omega)$ are both random variables. The autospectral density is written formally as

$$\Phi_p(\vec{x},\omega)\ \delta(\omega-\omega') = 2\pi \langle p(\vec{x},\omega)\, p^*(\vec{x},\omega)\rangle \tag{4}$$

where the asterisk denotes the complex conjugate and the brackets $\langle\rangle$ denote the formal ensemble average. The cross spectral density is

$$\Phi(\vec{x}_2,\vec{x}_1,\omega)\ \delta(\omega-\omega') = 2\pi \langle p(\vec{x}_2,\omega)\, p^*(\vec{x}_1,\omega)\rangle \tag{5}$$

so that the mean-square pressure is

$$\overline{p^2} = \langle p^2\rangle = \int_{-\infty}^{\infty} \Phi_p(\vec{x},\omega)\, d\omega = \lim_{\vec{x}_2\to\vec{x}_1} \int_{-\infty}^{\infty} \Phi_p(\vec{x}_2,\vec{x}_1,\omega)\, d\omega \tag{6}$$

In the monopole radiated sound, we note that the integral reduces to a particularly simple form in the case of a source volume that is small compared with an acoustic wavelength $2\pi/k_o$, i.e.

$$p_r(\vec{x},\omega) = \frac{\dot{Q}(\vec{y},\omega)}{4\pi r}\, e^{ik_o r} \tag{7}$$

where $\dot{Q}(\vec{y},\omega)$ is the rate of net mass injection into the acoustic space, and the magnitude of the radiated sound pressure is both independent of sound speed in the medium and independent of angular orientation of the source with respect to the observer.

Dipole sound is emitted when there is no net mass injection into the fluid, $\dot{q}=0$ but there is a distribution of forces oriented in the i-direction, say f_i per unit volume, then the corresponding expression for the radiated sound in free space (no reflectors) is

$$p_r(\vec{x},\omega) = \frac{-1}{4\pi} \int_{\text{SOURCE REGION}} f_i(\vec{y},\omega) \frac{\partial}{\partial y_i}\left(\frac{e^{ik_o r}}{r}\right) dV(\vec{y}) \tag{8}$$

If we assume that the dimensions of the source region are all smaller than $2\pi/k_o$ this dipole sound pressure component at frequency ω is

$$p_r(\vec{x},\omega) = \frac{-i\omega \cos\phi}{4\pi r c_o} |F_i| e^{ik_o r} \tag{9}$$

where F_i is the net force acting on the fluid in the i direction and ϕ is the angle to that direction as shown in Figure 3.

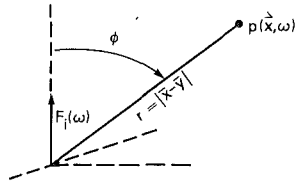


Figure 3 - Co-ordinate System of Dipole Source $F_i(\vec{y},\omega)$ and Field Point $p(\vec{x},\omega)$.

The directional nature of the force distribution accounts for the gradient of the free space Green function in the direction of the force and in turn the presence of k_o in equation (7). Comparing equations (7) and (9) two features are obvious; the appearance of the $\cos\phi$ directivity factor and the speed of sound in the denominator. Obviously, the more incompressible the fluid is (the larger C_o) the less sound will be radiated.

When the unbounded medium has no net mass injection and no net forces, the source distribution is representable as a system of force couples and shear stresses. The classical quadrupole[1,2] system may be expressed as

$$p_r(\vec{x},\omega) = \frac{1}{4\pi} \int_{\text{SOURCE REGION}} \tau_{ij} \frac{\partial^2}{\partial y_i \partial y_j}\left(\frac{e^{ik_o r}}{r}\right) dV(\vec{y}) \tag{10}$$

where τ_{ij} is the fluid stress tensor,

$$\tau_{ij} = \rho u_i u_j + (p - \rho c_o^2)\delta_{ij} - \tau'_{ij} \tag{11}$$

In this equation $\rho u_i u_j$ is the turbulent Reynolds stress for which the cross terms $i \neq j$ will vanish identically in irrotational flow. This term is intimately connected with the vorticity in the source region of turbulent flow. The term $(p-\rho c_o^2)$ expresses the deviation of pressure fluctuations in the source region from the isentropic pressure fluctuations in the acoustic medium outside the source region. This term becomes important for hot jets, non isentropic flow, and jets in which the speed of sound differs measurably from that in the external medium.

τ'_{ij} is a contribution from the viscous stresses and is generally neglected. The entire source term is assumed reduce to $\rho u_i u_j$ under certain limiting circumstances for which the sound speed and density in the shear flow equals that in the external flow, where $p = \rho c_o^2$ and, essentially, when the eddy viscosity $\nu_{ij} = |\overline{u_i u_j}| / dU_i/dy_j$ is much larger than the fluid kinematic viscosity.

To the extent that these relative magnitudes apply to free shear layers, such as jets (and they generally do for low mach number, single phase jets at high Reynolds number) it is conceptually quite possible to determine $p_r(\vec{x},\omega)$ as a function of the Fourier transform of $\tau_{ij} \simeq \rho u_i u_j$ given that the source distribution can be defined well enough to evaluate the integral. Unfortunately, as is well known, such a specification of $u_i u_j$ has been relatively elusive and for practical purposes equation (10) should be examined in a dimensional or in an instructional sense.

The same really holds true for other equations (7), (8) and (9) encountered in engineering practice. Much of the value of these simple results, and the theoretical work which has followed their initial presentation, lies in their usefulness in developing approaches to designing measurement evaluation programs, evaluating experimental data, and developing flow-acoustic similutude requirements. They also are widely used for designing simple experiments devised to test specific concepts of physical source characteristics. In fact only in certain relatively idealized circumstances is it possible to predict from fundamentals and first principles the sound levels that are generated in particular cases. Always, one must resort to carefully planned and executed measurement programs in order to provide quantitative evaluations of flow-induced sound and vibration. What stems from theories based on the above cannonical notions of acoustic sources are: the dependence of sound on specific case variables such as speed and geometric size; and the effects of sizes and shapes of boundaries contiguous to the flow on the process of conversion of kinetic energy of mean flow into sound waves. When surfaces are present, the free space Green function ($G(\vec{x}/\!/\vec{y},\omega)$) appearing in equations (1), (8), and (10) is replaced by a function $G(\vec{x}/\!/\vec{y},\omega)$ to obtain integrals of the type, e.g. dipole

$$p_r(\vec{x},\omega) = -\int \frac{\partial f_i(\vec{y},\omega)}{\partial y_i} G(\vec{x}/\!/\vec{y},\omega)\, dV(\vec{y})$$

$$= \int_{\substack{\text{SOURCE}\\ \text{REGION}}} f_i(\vec{y},\omega)\, \frac{\partial}{\partial y_i} G(\vec{x}/\!/\vec{y},\omega)\, dV(\vec{y}) \qquad (12)$$

$G(\vec{x}/\!/\vec{y},\omega)$ is the Green, or influence function with dimensionality (length) for a unit source at $\vec{y}$ generating sound at $\vec{x}$. The functions are solutions of the reduced wave equation simultaneously satisfying the particular boundary conditions for the surface under study. For example, as illustrated in Figure 4, for rigid boundaries, or boundaries on which the motion normal to the surface may be specified, the gradient of $G(\vec{x}/\!/\vec{y},\omega)$ in a direction normal to the surface is set to zero when $\vec{y}$ is on the surface.

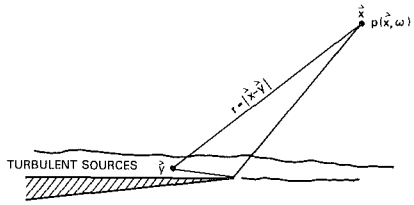


Figure 4 - Illustration of a Source Region Adjacent to an Accoustically Non-Compact Surface; in This Case a Wedge.

In these instances a second term must be added giving

$$p_r(\vec{x},\omega) = \int_{\substack{\text{SOURCE}\\ \text{REGION}}} \frac{\partial f_i(\vec{y},\omega)}{\partial y_i} G(\vec{x}/\!/\vec{y},\omega)\, dV(\vec{y}) + \int_{\text{SURFACE}} \frac{\partial p(\vec{y},\omega)}{\partial n} G(\vec{x}/\!/\vec{y},\omega)\, dS(\vec{y}) \qquad (13a)$$

where

$$\frac{\partial p(\vec{y},\omega)}{\partial n} = -i\omega\rho_o u_n(\vec{y},\omega) \qquad (13b)$$

and $u_n(\vec{y},\omega)$ will be zero on a rigid surface.

In some special instances, and high speed jet noise must be considered as one, the Green function $G(\vec{x}/\!/\vec{y},\omega)$ is dependent on the velocity of the flow. This is because the sound sources in the jet are moving relative to the ambient, still, fluid surrounding the jet. If U_c is the velocity of convection of the sound sources then $G(\vec{x}/\!/\vec{y},\omega)$ also depends on C_a where C_a is the speed of sound in the ambient fluid. The cause of this dependence is the refraction of sound waves at the interface between the moving jet fluid and the ambient fluid. At very low Mach number and cold jets this effect is minimal. Another change in the Green function will occur simply because of mean convection of the sources relative to the observer. This effect is common in sounds from high-velocity rotating machinery. A third complication arises when a mean vorticity, or shear layer occurs between the source region and the acoustic medium. In these cases the possibility exists for $G(\vec{x}/\!/\vec{y},\omega)$ to have instability modes in response to the source motions. Special cases related to this occur that invalidate the representation of equation (12). It is to be noted that equation (12) implies that flow sources drive the acoustic medium for which the behavior is governed by the Green function. There are cases in which the sound interacts with the fluid mechanics to cause reinforcements. Such cases are jet-edge interactions (in organ pipes and flutes), jet-hole interactions (in tea kettles and strainer plates) and wake-body interactions (in some laminar flow airfoils). These cases are often of practical significance, yet even a prediction of their frequency of occurrence (much less of the acoustic amplitudes experienced) is more

empirically than theoretically based. In a mathematical sense, one could say in equation (12) that both $G(\vec{x}//\vec{y},\omega)$ and $f_i(\vec{y},\omega)$ are dependent on the speed of sound in the medium and that G is also dependent on the value of $f_i(\vec{y},\omega)$.

3.0 General Principles of Similitude

The above equations are useful for writing down similitude relationships that may be used for scaling measurements in one case to obtain expected sound levels in another, similar, case. The similitude will be based on length, time, and force scales that are appropriate. Letting the appropriate length scale be L then the appropriate time scale is L/U where U is a measure of the fluid velocity in the machine.

The fluctuating forces that are generated in dynamically similar fluid machines will then be in the proportion.

$$F = gL^2 = \tfrac{1}{2}\rho_o U^2 L^2 \tag{14}$$

The sound pressure level at frequency $\omega = 2\pi f$ and in the frequency band $\Delta\omega = 2\pi\Delta f$ is defined as

$$L_s = 10 \log \overline{p^2}/p_o^2$$

where $p_o = 20\times10^{-6}\ n/m^2$ in aeroacoustics and $p_o = 10^{-6}\ n/m^2$ in hydroacoustic measurements. Letting

$$L_g = 20 \log g/p_o$$

then the following similitude relationships hold for the classical sound sources in free space or in the presence of acoustically compact or non reflective surfaces. Letting M_o, D_o and Q_o be particular universal spectrum functions of $\omega L/U$ we have for monopoles.

$$L_s = M_o + L_g + 20 \log L/r + 10 \log V_c/L^3 \tag{15a}$$

with a speed dependence at constant scaled frequency

$$L_s \sim 40 \log U \tag{15b}$$

For dipoles,

$$L_s = D_o + L_g + 20 \log U/c_o + 20 \log L/r + 10 \log A_c/L^2 \tag{16a}$$

with a speed dependence at constant scaled frequency

$$L_s \sim 60 \log U \tag{16b}$$

Finally for quadrupoles

$$L_s = Q_o + L_g + 40 \log U/c_o + 20 \log L/r + 10 \log V_c/L^2 \tag{17a}$$

with a speed dependencies at constant scaled frequency

$$L_s \sim 80 \log U \tag{17b}$$

The factors A_c and V_c are, respectively, the effective correlation area and volume of the sources; these are determined by the details of the integral combination of Green function and the source distribution in each particular circumstance. This type of combination will be made more clear below when we review the flow-excitation of a circular cylinder. Quite often, and particularly at very low Mach number, $U/c_o << 1$ these factors are closely represented by values of strictly hydrodynamic correlation areas and volumes that may be determined by aero (hydro) dynamic measurements of turbulence. In experiments which are scaled according to Mach number, $\left(\frac{U}{c_o}\right)_{MODEL} = \left(\frac{U}{c_o}\right)_{FULL\ SIZE}$ the relationship between the strictly hydrodynamic correlation area (or volume) and the

value determined by weighting the Green and source functions will be in nearly the same proprotion. It must always be kept in mind, however that in general the weighting of $G(\vec{x}/\!/\vec{y}_s, \omega)$ and the source function in this regard is necessarily frequency dependent and only in extremely simple circumstances (e.g. flow-dipole sound of rigid compact cylinders) is it a frequency-independent factor.

The functions M_o, D_o, and Q_o are strictly dimensionless and are all particular to the given circumstance or type of flow-body interaction, and they contain all the frequency dependence that is controlled by the flow. Accordingly, for example, D_o may be representable as a spectrum function multiplied by a frequency band, i.e., starting with a basic definition of the factor

$$D_o = 10 \log \left\{ \int_{\frac{L\Delta\omega}{U}} \phi\left(\frac{\omega L}{U}\right) d\left(\frac{\omega L}{U}\right) \right\} \tag{18a}$$

it may be expressed as a dimensionless spectral density evaluated at $\omega = \omega_o$ multiplied by a bandwidth $\Delta\omega$ i.e.

$$D_o \simeq 10 \log \left\{ \phi\left(\frac{\omega_o L}{U}\right) \frac{L\Delta\omega}{U} \right\} \tag{18b}$$

as long as $\phi\left(\frac{\omega_o L}{U}\right)$ is broadband. If $\phi\left(\frac{\omega_o L}{U}\right)$ is tonal, then

$$D_o \simeq D_o\left(\frac{\omega L}{U} = 2\pi S\right) \tag{18c}$$
$$= 0, \quad \frac{\omega L}{U} \neq 2\pi S$$

where S is a Strouhal number of the flow.

Similar notions hold for the other factors M_o and Q_o.

4.0 Correlation Areas and Fluid - Structure Coupling

More needs to be said of the factors A_c/L^2 and V_c/L^2 and to do so requires some manipulation of the integral relationships. For a one-dimensional structure as a taut wave these ratios are replaced by $2\Lambda_3/L$ where $2\Lambda_3$ is the correlation length of the flow disturbance along the length. Taking equation (8) as an example, and representing it in terms of a surface pressure distribution on a rigid surface as

where $S(\vec{y}_s)$ describes the surface and $n_i(\vec{y}_s)$ is the direction cosine in the i direction, then

$$p_r(\vec{x}, \omega) = \int_0^{2\pi} \int_0^{L} p(\vec{y}_s, \omega)\, n_i \frac{\partial g(\vec{x}/\!/\vec{y}_s, \omega)}{\partial n}\, dS(\vec{y}_s) \tag{19}$$

in the limit of $r = |\vec{x} - \vec{y}_s| >> |\vec{y}_s|$ and $\omega |\vec{y}_s| / c_o << 1$, and where

$$g(\vec{x}/\!/\vec{y}_s, \omega) = \frac{1}{4\pi r} \exp(ik_o r)$$

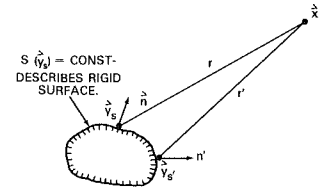


Figure 5 - Geometry of a Pressure Distribution on a Surface S(y,) = Const. on Which the Unsteady Flow-Induced Pressure is Described.

is the free-space Green function. Such a situation is depicted in Figure 5. We assume that the function $p(\vec{y}_s, \omega)$ can be measured and it may be represented as a cross-spectral density of pressures (see also equations (4) through (6)) at $\vec{y}_s$ and $\vec{y}_s'$; accordingly using equations

$$\Phi_p(\vec{y}_s, \vec{y}_s', \omega)\, \delta(\omega - \omega') = 2\pi \langle p(\vec{y}_s', \omega)\, p^*(\vec{y}_s, \omega) \rangle$$

represent this cross spectral density then we have for the spectral density of the radiated sound pressure at a point

$$\Phi_{p_r}(\vec{x},\omega) = \int_S \int_{S'} \Phi_p(\vec{y}_s,\vec{y}_s',\omega)\, n_i(\vec{y}_s)\, n_i(\vec{y}_s') \left(\frac{\partial g}{\partial n}\right)^* \left(\frac{\partial g'}{\partial n'}\right) dS(\vec{y}_s')\, dS(\vec{y}_s) \tag{20}$$

n' and g' denote values at $\vec{y}_s'$. The values of $g(\vec{x}/\vec{y},\omega)$ and $g(\vec{x}/\vec{y}',\omega)$ are coupled together if the cross spectrum junction does not separate into a pair of functions of the form

$$\Phi_p(\omega)\, A(\vec{y}_s)\, A(\vec{y}_s')$$

Since products of g and g' may be represented as functions of

$$\omega(|\vec{y}_s| - |\vec{y}_s'|)/c_o = k_o(|\vec{y}_s| - |\vec{y}_s'|)$$

the degree of coupling will depend on the value of k_o in relation to the spatial scale of the cross spectrum of

$$\Phi_p(\vec{y}_s,\vec{y}_s',\omega)$$

Further simplification is generally necessary. If we say that the cross spectrum function is independent of the end points or extremities of the flow, on the surface and that within the flow the sources are <u>spatially homogeneous</u> then we may write the correlation in terms of the difference function

$$\Phi_p(\vec{y}_s,\vec{y}_s',\omega) = \Phi_p(\vec{y}_s - \vec{y}_s',\omega)$$

That is, the spatial correlation depends only the separation $\vec{y}_s - \vec{y}_s'$ but the source level depends on the location $\vec{y}_s$ and frequency, ω. Noting the limiting process expressed by equation (6) a further simplification that is generally made is to define a correlation function by

$$R(\vec{y}_s' - \vec{y}_s,\omega) = \frac{\Phi(\vec{y}_s' - \vec{y}_s,\vec{y}_s,\omega)}{\Phi(\vec{y}_s,\omega)} \tag{21}$$

and a <u>correlation area</u> by

$$A_c(\omega) = \iint_{-\infty}^{\infty} R(\vec{\xi}_s,\omega)\, d^2\vec{\xi}_s \tag{22}$$

This function, so defined, depends only on a knowledge of the fluid mechanics and not on any attention given to the acoustics of the problem.

Equation (20) may now be put into perspective since we may write

$$\Phi_{p_r}(\vec{x},\omega) = \int_S \Phi_p(\vec{y}_s,\omega)\, n_i(\vec{y}_s) \cdot \left\{ \int_s R(\vec{\xi}_s,\omega)\, n_i(\vec{y}_s + \vec{\xi}_s) \left[\frac{\partial g^*(\vec{y}_s,\omega)}{\partial n}\right] \left[\frac{\partial g(\vec{y}_s + \vec{\xi}_s,\omega)}{\partial n}\right] d^2\vec{\xi}_s \right\} d^2\vec{y}_s \tag{23}$$

What was represented above in equation (16a) as a correlation area A_c is really the entire term in curley brackets. In many problems of flow-induced noise simplification of this term requires a good deal of assumption regarding the acoustics-flow interaction. Often, however, if $A_c << (2\pi/k_o)^2$ equation (23) may be reduced to

$$\Phi_{p_r}(\vec{x},\omega) \simeq A_c(\omega) \int_S \Phi_p(\vec{y}_s,\omega)\, n_i^2(\vec{y}_s) \left| \frac{\partial g(\vec{y}_s,\omega)}{\partial n} \right|^2 d^2\vec{y}_s \tag{24}$$

For example, in the case of a flow past a rigid cylinder the surface pressure distribution replacing the one above gives a new source strength

$$f_i(\vec{y},\omega) = p(a,\phi_o,z_o,\omega)\, \delta(r-a)$$

where a=d/2. Furthermore $\partial/\partial n = \partial/\partial r$ on the surface of the cylinder and $n_i = \cos\phi$ for the co-ordinate systems shown in Figures 1 and 3. Also, since the surface pressure $p(a,\phi_o,z_o,\omega)$ due to vortex shedding is deterministic in ϕ_o but random in z_o the cross spectral density is of the form

$$\Phi_p(\vec{y}_s,\vec{y}_s',\omega) = p_o^2(\omega) \cos\phi_o \cos\phi_o' R(z_o' - z_o)$$

where $R(z_o' - z_o)$ expresses the axial correlation, p_o is the amplitude of pressure at a point on the upper or lower surface $\phi_o = o, \pi$. The lift force per unit length illustrated in Figure 1 is

$$f_1'(z_o,t) = \int_0^{2\pi} p(a, \phi_o, z_o, \omega) \cos\phi_o \, a \, d\phi_o$$

which may be written in the form

$$f_1'(z_o,t) = \pi a p_o e^{i\alpha(z_o)}$$

if the pressure is

$$p(a, \phi_o, z_o, \omega) = p_o(\omega) \cos\phi_o \, e^{i\alpha(z_o)}$$

where $\alpha(z_o)$ is a random phase function of axial location. This phase function is

$$\langle e^{i\alpha(z_o')} \, e^{-i\alpha(z_o)} \rangle = R(z_o' - z_o)$$

5.0 The Formal Description of Flow-Induced Vibration

Many problems of flow-induced sound involve surface vibration, and this is governed in a general way by relationships of the type in equation (13a). Following the above we can write down the spectrum of radiated sound at a point that is far from the vibrating body:

$$\left[\Phi_{p_r}(\vec{x},\omega)\right]_{VIB} = \omega^2 \rho_o^2 \int_S \int_{S'} \Phi_{u_n}(\vec{y}_s', \vec{y}_s, \omega) \cdot G^*(\vec{x}\,//\,\vec{y}_s, \omega) \, G(\vec{x}\,//\,\vec{y}_s', \omega) \, d^2\vec{y}_s \, d^2\vec{y}_s' \tag{25}$$

where $G(\vec{x}//\vec{y}, \omega)$ is the Green function for the particular surface geometry in question and where $\Phi_{u_n}(\vec{y}_s', \vec{y}_s, \omega)$ is the cross-spectral density of surface velocity normal to the undisturbed surface. The assumption is made that the deformation amplitude of the surface is very small. At this point we assume that the surface may be described as a summation of a discrete set of normal modes, each with a distinct geometric pattern representable with dimensionless, orthogonal, modal shape functions $\Psi_{mn}(\vec{y}_s)$ that are normalized so that

$$\int_S \Psi_{mn}(\vec{y}_s) \, \Psi_{pq}(\vec{y}_s) \, d^2\vec{y}_s = A_s \, \delta_{mnpq} \tag{26}$$

where $\delta_{mnpq} = 1$ for $m = p$ and $n = q$ and otherwise, and where A_S is the area of the surface. The presence of δ_{mnpq} in equation (26) expresses the fact that if we find a way to purely excite mode mn we will not excite the other modes. This makes all modes decoupled and uncorrelated so that the total kinetic energy of the surface is just a summation of the kinetic energies of the vibration modes of the surface.

In the special case of a one dimensional structure, we let the indices m and n represent an ordering of modes, the mode m vibrates at frequency ω_m mode n at frequency ω_n . These frequencies are determined by the geometry of the surface and the wave speed of waves which maybe induced in that material if it were unbounded. As an example of a one dimensional structure an effectively infinite taut wire[7] with mass per unit length m_c and under tension T may sustain bending waves which travel along it at speed

$$c = \sqrt{\frac{T}{m_c}}$$

A finite taut string of length L will have resonance frequencies

$$\omega_n = n\pi c/L \qquad \text{where} \qquad n = 0, 1, 2, \cdots \text{etc}$$

and a mode shape

$$\Psi_n(y) = \frac{1}{\sqrt{2}} \sin(n\pi z_o/L) \qquad \text{for} \quad o \le z_o \le L \tag{27}$$

The average frequency interval between resonances for modes n and $n+1$ will be $\Delta\omega_n = \pi c/L$, or the number of modes per unit frequency $n(\omega)$ will be

$$n(\omega) = L/\pi c$$

Thus described, the vibration velocity is often assumed to be representable as a second order oscillator where each value of resonance frequency indicates an independent degree of freedom. Therefore, the cross spectral density of velocity of any two dimensional surface can generally be expressed as a summation.

$$\Phi_{u_n}(\vec{y}_s, \vec{y}_s', \omega) = \sum_{mn} \frac{\omega^2 \psi_{mn}(\vec{y}_s)\, \psi_{mn}(\vec{y}_s')}{M^2 \omega_{mn}^4 |Z|^2} \cdot \tag{28}$$

$$\left\{ \int_S \Phi_p(\vec{z}_s, \omega) \int_S R(\vec{\xi}, \omega)\, \psi_{mn}(\vec{z}_s)\, \psi_{mn}(\vec{z}_s + \vec{\xi})\, d^2\vec{\xi}\, d^2\vec{z}_s \right\}$$

where the cross spectrum function is the same as that appearing in equation 20, and where

$$|Z|^2 = \left[1 - \left(\frac{\omega}{\omega_{mn}}\right)^2 \right]^2 + \eta^2 \left(\frac{\omega}{\omega_{mn}}\right)^2 \tag{29}$$

is an impedance function. M is the mass and η is the loss factor of the mode which is equal to the log decrement divided by π. The terms outside the one in curley brackets for a given frequency ω_{mn} emerge simply as the velocity response of the nth mode to a unit force of excitation. The term in curley brackets represents the effective mean-square force that excites the surface into the particular 2-dimensional mode which we have designated by mn. The effective input force depends on how the spectrum function $\Phi_p(\vec{y}_s, \vec{y}_s', \omega)$ matches the mode shape functions $\psi_{mn}(\vec{y}_s)$. All the qualifications regarding the relationship between the flow and the Green function that were discussed in connection with equation (23) above apply here to the mode shape function. In the case of structural response, the relationship between the flow and how the structure may accept energy from the flow depends in part on how A_c (equation (22)) compares with the square of the structural wavelength, $(2\pi\omega/c_b)^2$ where c_b is the speed of free flexural waves on the surface. If $A_c < (2\pi\omega/c_b)^2$ then analogously to equation 24 we have

$$\Phi_{u_n}(\vec{y}_s, \vec{y}_s', \omega) \simeq \sum_{mn} \frac{\omega^2 \psi_{mn}(\vec{y}_s)\, \psi_{mn}(\vec{y}_s')}{M^2 \omega_{mn}^4 |Z|^2} \int_S \Phi_p(\vec{z}_s, \omega)\, A_c(\omega)\, |\psi_{mn}(\vec{z}_s)|^2\, d^2\vec{z}_s \tag{30}$$

where the integral may often be effectively written as

$$\overline{\Phi_p(\omega)}\; A_s\, A_c(\omega) \tag{31}$$

and $\overline{\Phi_p(\omega)}$ is the spectrum of excitation pressure averaged over the surface.

Equation (30) is then substituted into equation (25) to evaluate the acoustic field. The sound pressure depends on how $\psi_{mn}(\vec{y})$ matches $G(\vec{x}/\!/\vec{y}, \omega)$ for the particular geometry being considered. Since $\psi_{mn}(\vec{y})$ are prescribed mode shapes that describe the surface motion and since $G(\vec{x}/\!/\vec{y}, \omega)$ is determined for the class of acoustic problems involving that surface, the integral of the product of $G(\vec{x}/\!/\vec{y}, \omega)$ and $\psi_{mn}(\vec{y})$ over the surface is determined independently of the statistical nature of flow excitation.

One complication arises, however, when the coupling of $R(\vec{y}_s' - \vec{y}_s)$ and $\psi_{mn}(\vec{y}_s)$ and of $G(\vec{x}/\!/\vec{y}, \omega)$ and $\psi_{mn}(\vec{y})$ are dominated by different sets of modes. In such cases the vibration and acoustic fields may be controlled by different modes. This case may arise in flow-induced vibration of continued structures such as aircraft fuselages. In such cases the structural modes that are resonant at frequency are best excited by the flow length scale U/ω. Those modes which make the most sound couple well with scales c_o/ω. Accordingly if we are interested in flow-induced sound, we must first identify those mode orders, say p, q with a maximum value of

$$\int_{A_s} G(\vec{x}/\!/\vec{y}_s, \omega)\, \psi_{p,q}(\vec{y}_s)\, d^2\vec{y}_s \tag{32}$$

prior to evaluating bracketed term of equation (28).

A second complication arises when evaluating the damping, or loss factor. This factor includes the effects of all mechanisms of power flow from the structure: hysteretic material damping, applied damping, damping due to structural coupling with adjacent (or internal) structures, apparent damping due to acoustic energy radiated away from the structure, damping due to viscosity in the adjacent fluid, and aero (hydro) - dynamic damping due to interaction between the surface and the fluid motion (e.g., due to lifting surface motion). The net response of all modes that are resonant in a given frequency band is therefore controlled by the damping of the individual modes. Therefore in carrying out the summation in equation (28) one must exercise care in making approximations that one or more terms in the summation are constants for the values of η that are appropriate.

A third complication arises when there is an aero (hydro) - elastic coupling between the fluid and the surface. In such cases $\Phi_p(\vec{y}_s, \vec{y}_s', \omega)$ may be dependent on u_n in such a manner that the flow and structure are deterministic. In these cases even the loss factor η must be adjusted (generally downward) to account for the phased injection of energy from the flow into the structure (generally a bluff body with a discrete trailing edge flow). This phenomenon is related to singing hydrofoils (and propellers), self-excitation in tube bundles.

It is also important to acknowledge that the damping measured, say as commonly done by reverberation times without flow, may be controlled lightly damped modes which do not radiate much sound (therefore have low acoustic damping) or they may become more heavily damped due to aerodyanmic damping when the flow is turned on. Accordingly reverberation time measurements may be misleading when used to deduce loss factors of a structure.

6.0 Simplitude For Sound From Flow Induced Vibration.

Using equation (30) we note that the flow-induced acceleration,

$$\overline{a^2} = \int_{\Delta\omega} \omega^2 \Phi_{u_n}(\vec{y}_s, \omega)\, d\omega = \int_{\Delta\omega} \Phi_a(\vec{y}_s, \omega)\, d\omega$$

may be expressed in dimensional form as

$$\overline{a^2}(\vec{y}_s) = \left[\begin{array}{c}\text{DIMENSIONLESS}\\ \text{SPECTRUM}\\ \text{FUNCTION}\end{array}\right] \cdot \frac{g^2 A_c A_s}{M^2 \eta} \tag{32}$$

or, referring the acceleration to $1g = 980\ cm/sec^2$

and letting $L_a = 10 \log \overline{a^2}/g^2$

we have the acceleration level of the mn^{th} mode

$$L_{a_{mn}} = a_o + L_g + 10 \log A_c/A_s - 20 \log \rho_p \omega_{mn} - 10 \log \eta \tag{33}$$

where A_c is the effective correlation area for flow-structure coupling, and where A_o is a spectrum function along the lines of M_o, D_o, and Q_o of equation (15), (16), and (17). The structure-acoustic coupling is determined from equation (25) by noting that the sound pressure averaged over a spherical surface of area $4\pi r^2$ where $|\vec{x}| = r$ surrounding the source region is given by

$$\overline{\Phi_{p_r}(\omega)} = \frac{1}{4\pi r^2} \int_{4\pi r^2} \Phi_{p_r}(\vec{x}, \omega)\, d^2\vec{x}$$

$$= \rho_o^2 c_o^2 \frac{A_s}{4\pi r^2} \sigma_{rad} \langle v^2 \rangle \tag{34}$$

where the radiation efficiency[8,9] is defined as

$$\frac{A_s}{4\pi r^2} \sigma_{rad} = \frac{1}{4\pi} \int_{\Omega=4\pi} k_o^2 A_s \left[\int_S \psi_{mn}(\vec{y}_s)\, G(\vec{x}/\!/\vec{y}_s, \omega) \frac{dS(\vec{y}_s)}{A_s} \right]^2 d\Omega \tag{35}$$

The radiation efficiencies of a variety of planar and non planar structures are available in the literature.

Combining equations (32) and (33) we find the similarity rule for the sound pressure level radiated by all the resonant modes at in the frequency band $\Delta\omega$ is

$$L_s = a_o + L_q + 20\log\left(\frac{\rho_o c_o}{\rho_p h \omega}\right) + 10\log\frac{A_c}{L^2} + 20\log\frac{L}{r} + 10\log\sigma_{rad} + 10\log\eta_s(\omega)\Delta\omega \qquad (36)$$

The derivation of equation (36) presumes that the radiation from each of all the resonant modes in band width $\Delta\omega$ is nearly the same so that the net sound power level is just the representative level in the band multiplied by the number of resonant modes in the band.

For flat surfaces $\eta_s(\omega) = \sqrt{3}\, A_s\, (2\pi h C_L)^{-1}$

where C_l is the longitudinal wave speed in the plate; $C_b = \sqrt{\omega h C_L/\sqrt{12}}$

7.0 References

1. Cramer, H. and Ledbedder, M.R. "Stationary and Related Stochastic Processes" J. Wiley & Son, 1967

2. Lighthill, J. "On Sound Generated Aerodynamically I, General Theory". Proc. Roy Soc. 211A, 564-587, 1952

3. Lighthill, J. "On Sound Generated Aerodynamically II, Turbulence as a Source of Sound" Proc. Roy. Soc. 222A, 1-32, 1954

4. Curle, N. "The Influence of Solid Boundaries upon Aerodynamic Sound" Proc. Roy. Soc. A231, 505-514, 1955

5. Crighton, D.G. "Basic Principles of Aerodynamic Noise Generation" Prog, Aerospace Sci. 16 131-96, 1975

6. Lin, Y.K. "Probabilistic Theory of Structural Dynamics" McGraw-Hill, 1967

7. Morse, G. "Vibration and Sound ", McGraw-Hill, 1948

8. Maidanik G. "Response of Ribbed Panels To Reverberant Acoustic Fields", J. Acoust. Soc. Am., 34, 809-826, 1962

9. Lyon, R.H. "Statistical Energy Analysis of Dynamical Systems: Theory and Applications" M.I.T. Press, 1975

EXTENDED SOUND SOURCES

M. C. Junger
Cambridge Acoustical Associates, Inc.
54 Rindge Avenue Extension
Cambridge, Massachusetts 02140, USA

The point source solution is used to construct the far-field of line-arrays and its extension to planar sources is described. The alternative, more generally applicable transform technique, whereby the far-field is evaluated analytically by the method of stationary-phase, is derived for axisymmetric planar radiators and illustrated for rigid circular piston sources. The lecture concludes with the transform solution of the sound field radiated by a finite cylinder of arbitrary dynamic configuration.

So far we have considered one-dimensional sound fields in the form of plane waves and spherical waves radiated by uniformly pulsating spherical radiators and point sources. The latter will be used to construct extended sources.

A. LINE ARRAYS

1. Synthesis from Point Sources

The line array can be used to illustrate the characteristics of more complicated planar and convex sources. The sound field of a line array coinciding with the x-axis can be synthesized by considering it as a juxtaposition of point sources of volume acceleration

$$\frac{d\ddot{Q}}{dx}\,dx\,\exp[i\gamma(x)] \quad , \quad -L < x < L$$

The far-field contribution of an array element located at a point x is constructed from Eq. 26 of Lecture 5, identified here as Eq. 5.26. Allowance must be made for the fact that the radius vector $\bar{R}$ is no longer measured from the point source but from some other point, say the midpoint of the array, x=0. The distance R in Eq. 5.26 is therefore replaced by $|\bar{R}-\bar{R}_o|$, where $\bar{R}_o$ defines the location of an array element:

$$p_s(|R-R_o|) = \frac{\rho\ddot{Q}}{4\pi|R-R_o|}\,\exp(ik|R-R_o|) \tag{1}$$

Clearly the sound field is symmetrical about the x-axis, but it will be a function not only of R but also of the spherical angle θ between the x-axis and the radius vector R. Using elementary trigonometry to express $|R-R_o|$ in terms of R, x, and θ, and integrating Eq. 1 over the length $-L < x < L$ of the array, the sound field becomes:

$$p(R,\theta) = \frac{\rho}{4\pi}\int_{-L}^{L}\frac{(d\ddot{Q}/dx)\exp[i\gamma(x)+ik(R^2+x^2-2Rx\cos\theta)^{1/2}]}{(R^2+x^2-2Rx\cos\theta)^{1/2}}\,dx$$

Unless $R \gg x$, i.e., $R \gg L$, and $R \gg kL^2$, this θ-dependence is a function of R, thus violating one of the far-field criteria. Another far-field criterion, viz. that the pressure vary as R^{-1}, is similarly violated if this condition is not fulfilled.

2. The Far-Field

When the far-field requirements are satisfied, one can set

$$(R^2+x^2-2Rx\cos\theta)^{1/2} \simeq R-x\cos\theta+0(x^2/R)$$

Note that while the x-term in the denominator can be dropped, the phase angle kx must be retained no matter how large kR. The integral specialized to the far-field now becomes

$$p(R,\theta) = \frac{\rho e^{ikR}}{4\pi R}\int_{L}^{L}\frac{d\ddot{Q}}{dx}\exp(i\gamma-ikx\cos\theta)\,dx \quad , \quad x/R \ll 1 \quad , \quad kx^2/R \ll 1 \tag{2}$$

If $d\ddot{Q}/dx$ is constant, and the phase shift a linear function of x, viz. $\gamma = k_o x$, the integral can be evaluated analytically:

$$p(R,\theta) = p_s(R)\,\frac{\sin(k_o-k\cos\theta)L}{(k_o-k\cos\theta)L} \tag{3}$$

where p_s is given in Eq. 1 with $\ddot{Q} = 2LdQ/dx$, the resultant volume acceleration of all the transducers. This peak pressure p_s occurs on the conical surface when the argument of the sine vanishes:

$$\theta_o = \cos^{-1} k_o/k \quad , \quad k_o < k$$

When $k_o=k$, we have an end-fire array whose peak lies on the array axis. If $k_o=0$, we have a broad-side array whose peak pressure occurs in the plane of symmetry $\theta = \pi/2$. The peak pressure drops when $k_o > k$, the array can be steered electrically by adjusting the phase shift γ, effectively delaying the electrical signal to successive array elements. This linear delay simulates a constant phase velocity ω/k_o which can also be implemented in practice with a mechanical waveguide. The larger kL, the more rapidly the pressure drops off as θ deviated from θ_o. If the array length is sufficiently large in terms of wavelengths, nulls and minor peaks appear in the directivity pattern. For example, for the broad-side array, nulls appear whenever $kL\cos\theta = n\pi$, where n is an integer. The array must measure at least one wavelength before this equation admits a solution. The increased sharpness of the pressure peak and the enhanced complexity of the directivity pattern as the ratio of characteristic source dimension to wavelength increases will be shown to apply to two-dimensional sources as well.

B. PLANAR SOURCES

1. Rayleigh's Formula

Two-dimensional planar arrays vibrating about a plane of symmetry coinciding with the midplane of the array elements can be modelled as a surface distribution of point sources whose contributions to the sound field are given by Eq. 1. The array thickness is assumed small in terms of wavelengths. The plane of symmetry is equivalent to a rigid baffle, since $\partial p/\partial n$ and hence $\ddot{w}$ vanishes over this plane. If S is the surface on one side of the plane of symmetry, the volume acceleration of an array element is $2\ddot{w}dS$. The denominator of the integrand therefore displays the factor 2 instead of 4, as in Eq. 1:

$$p(\bar{R}) = \frac{\rho}{2\pi} \iint_{S(\bar{R}_o)} \frac{\ddot{w}(\bar{R}_o)\exp(ik|\bar{R}-\bar{R}_o|)}{|\bar{R}-\bar{R}_o|} \, dS(\bar{R}_o) \tag{4}$$

If R_o/R, $kR_o^2/R \ll 1$, this integral simplifies to a far-field integral expression, effectively the extension of Eq. 2 to two dimensions. This is Rayleigh's Formula. We shall, however use a different approach which, for planar sources, leads to an equivalent integral expression in the form of an inverse transform. Our preference for the transform formulation is that it is applicable to non-planar sources such as the cylinder and to structure-fluid interaction problems. Both of these situations will be illustrated. First, however, the transform technique will be applied to an elementary planar source which could have been analyzed in terms of Eq. 4.

2. The Integral Transform Formulation of Axisymmetric Pressure Fields

In this approach, the wave equation and the boundary conditions are integral-transformed. After some simple algebra, the integral transform of the pressure field is solved for. Finally, the inverse transform of this result yields an integral expression for the actual pressure. This integral can be evaluated analytically in the far-field. Consider for example an axisymmetric pressure field. The Helmholtz equation in cylindrical coordinates is

$$(\nabla_s^2 + \frac{\partial^2}{\partial z^2} + k^2)p(r,z) = 0 \tag{5}$$

where the surface Laplace operator is

$$\nabla_s^2 \equiv \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} \tag{6}$$

For axisymmetric configurations, the most convenient integral transform is the Hankel transform of order zero:[1]

$$\tilde{f}(\gamma) = \int_0^\infty f(r)J_o(\gamma r)r\,dr$$

$$f(r) = \int_0^\infty \tilde{f}(\gamma)J_o(\gamma r)\gamma\,d\gamma \tag{7}$$

Applying this transform to the Helmholtz equation Eq. 5, and noting that[2]

$$\int_0^\infty \nabla_s^2 f(r)J_o(\gamma r)r\,dr = -\gamma^2\tilde{f}(\gamma)$$

the transformed Helmholtz equation becomes an ordinary differential equation which governs the pressure transform:

$$\left[\frac{d^2}{dz^2} + (k^2-\gamma^2)\right] \tilde{p}(\gamma;z) = 0 \tag{8}$$

This is in the form of the Helmholtz equation governing plane waves, Eq. 5.13. The general solution for waves propagating in the positive z-direction is

$$\tilde{p}(\gamma;z) = A\exp[i(k^2-\gamma^2)^{1/2}z] \tag{9}$$

The undetermined coefficient is computed from the transform of the boundary condition, Eq. 5.16:

$$\frac{\partial\tilde{p}}{\partial z} = -\rho\ddot{\tilde{w}}(\gamma) \quad , \quad z = 0 \tag{10}$$

Differentiating Eq. 9 and substituting the derivative in Eq. 10, the coefficient A can be solved for and substituted in Eq. 9:

$$\tilde{p}(\gamma;z) = \frac{i\ddot{\tilde{w}}(\gamma)\rho\exp[i(k^2-\gamma^2)^{1/2}z]}{(k^2-\gamma^2)^{1/2}} \tag{10a}$$

The formal expression for the pressure field is obtained by substituting Eq. 11 in the latter of Eqs. 7:

$$p(r,z) = i\rho \int_0^\infty \frac{\ddot{\tilde{w}}(\gamma)}{(k^2-\gamma^2)^{1/2}} J_o(\gamma r)\exp[i(k^2-\gamma^2)^{1/2}z]\gamma d\gamma \tag{11}$$

This result is applicable both in the near- and far-field. However, in the near-field the integral can only be evaluated at restricted locations, specifically on the axis where $J_o=1$. The radiation loading can be calculated by setting z=0, evaluating the integral $2\pi J_o(\gamma r)rdr$ over the piston surface, which yields an analytically integrable inverse transform for the force exerted by the fluid. Here we are exclusively concerned with the far-field.

3. The Point of Stationary-Phase and its Physical Interpretation

To obtain an expression tractable by asymptotic integration the Bessel function in Eq. 12 is expressed in terms of Hankel functions:

$$J_o(x) = \frac{1}{2}H_o^{(1)}(x) + \frac{1}{2}H_o^{(2)}(x)$$

$$= \frac{1}{2}H_o^{(1)}(x) - \frac{1}{2}H_o^{(1)}(-x)$$

The integral of $-H_o^{(1)}(-x)$ from 0 to ∞ equals the integral of $-H_o^{(1)}(x)$ from 0 to $-\infty$, i.e., or, alternatively, of $H_o^{(1)}(x)$ from $-\infty$ to 0. We thus finally have an integral from $-\infty$ to ∞. Specializing to the far-field, the asymptotic large-argument expression for the Hankel function is used:

$$H_n^{(1)}(x) \simeq \left(\frac{2}{\pi x}\right)^{1/2}(-i)^n\exp\left(ix - i\frac{\pi}{4}\right) \quad , \quad x >> n^2+1 \tag{12}$$

In the far-field, it is convenient to use spherical coordinates. Setting

$$r = R\sin\theta$$

$$z = R\cos\theta$$

Eq. 11 finally becomes

$$p(R,\theta) = \frac{\rho e^{-i\pi/4}}{(2\pi R\sin\theta)^{1/2}} \int_{-\infty}^{\infty} \frac{\ddot{\tilde{w}}(\gamma)\gamma^{1/2}}{(k^2-\gamma^2)^{1/2}} \times \exp\{iR[\gamma\sin\theta+(k^2-\gamma^2)^{1/2}\cos\theta]\}d\gamma \tag{13}$$

The integrand is of the form $F(\gamma)e^{if(\gamma)}$. For large R, the phase angle varies rapidly with γ compared to the modulus of the integrand. Consequently, regions of opposite phase tend to cancel, the main contribution coming from the point of stationary phase

$$df(\gamma)/d\gamma = 0 \quad , \quad \gamma = \bar{\gamma}$$

For the phase angle in Eq. 13,

$$\bar{\gamma} = k\sin\theta \tag{14}$$

The physical interpretation of this result can be envisioned by considering that the integral representation of the transform of the boundary configuration $\ddot{w}(r)$ encompasses the entire wavenumber spectrum. Only wavenumbers in the vicinity of the point of stationary-phase, Eq. 14, contribute to the pressure at field points located on the cone of vertex angle θ. It can be verified that $\lambda/\sin\theta$ is the trace wavelength, in the z=0

plane, of planar sound waves travelling in the θ-direction. Recalling that $k = 2\pi/\lambda$, it is apparent that the point of stationary-phase is associated with the trace wavenumber. The same interpretation applies to the cone θ_o where the peak pressure radiated by the line array is observed.

4. Asymptotic Stationary-Phase Evaluation of the Inverse Transform

Integrals in the form

$$I = \int_{-\infty}^{\infty} F(\gamma) e^{if(\gamma)}\, d\gamma \tag{15}$$

with F a slowly varying function of γ, can be evaluated symptotically by considering only the contribution of the point of stationary-phase. The result is[3]

$$I = F\left(\frac{2\pi}{\left|\partial^2 f/\partial\gamma^2\right|}\right)^{1/2} \exp(-i\,\frac{\pi}{4} + if) \quad , \quad \gamma = \bar{\gamma} \tag{16}$$

Substituting the function F and f from Eq. 13, with $\bar{\gamma}$ as given in Eq. 14, and noting that,

$$\left.\frac{\partial^2 f}{\partial\gamma^2}\right|_{\gamma=\bar{\gamma}} = -\frac{R}{k\cos^2\theta}$$

one obtains a concise expression for the far-field without restricting the dynamic configuration $\ddot{w}(r)$ of the radiating surface:

$$p(R,\theta) = \rho\ddot{\tilde{w}}(k\sin\theta)\,\frac{e^{ikR}}{R} \tag{17}$$

5. The Circular Piston

This result will now be applied to an idealized sound source simulating various practical sources such as a loudspeaker in a planar baffle or a cylinder, short in terms of wavelengths, vibrating about its plane symmetry. If the piston is rigid, motion is uniform over a circle of radius a:

$$\begin{aligned} \ddot{w}(r) &= \ddot{W} \quad , \quad r < a \\ &= 0 \quad , \quad r > a \end{aligned} \tag{18}$$

The Hankel transform of this function is

$$\begin{aligned} \ddot{\tilde{w}}(\gamma) &= \ddot{W}\int_0^a J_o(\gamma r)\, r\, dr \\ &= \frac{\ddot{W} a J_1(\gamma a)}{\gamma} \end{aligned} \tag{19}$$

Setting γ equal to its stationary-phase value, Eq. 14, and substituting the result in Eq. 17, one obtains the far-field

$$p(R,\theta) = \frac{\rho\ddot{W}ae^{ikR}}{R}\,\frac{J_1(ka\sin\theta)}{ka\sin\theta} \tag{20}$$

If we now let $ka \to 0$, $J_1(ka\sin\theta) \to ka\sin\theta/2$, and set $dS = \pi a^2$, we retrieve Rayleigh's formula specialized to a small source. If $ka > 3.8$, nulls appear. Also, the larger ka, the narrower the on-axis peak.

Time does not permit the analysis of rectangular sources. Since their dynamic configuration depends on two coordinates, the solution requires a double Fourier transform applied, as above, to both the Helmholtz equation in rectangular coordinates and to the boundary condition. The Fourier transfo m solution will now be illustrated for an important convex source.

C. CONVEX SOUND SOURCES

1. Mathematical Model of the Radially Vibrating Cylinder

The finite cylindrical radiator is not analytically tractable. It can, however, be approximated by a mathematical model which yields realistic analytical results except near the cylindrical axis. This model consists in extending the actual cylinder source with two semi-infinite coaxial cylindrical baffles of the same diameter. The problem is analyzed by means of transforms in a manner paralleling the solution of planar sources. The dynamic configuration of the actual source is expanded in a Fourier series

$$\ddot{w}(\phi,z) = \sum_{n=0}^{\infty} \ddot{W}_n(z)\cos n\phi \tag{21}$$

The pressure field is similarly expanded

$$p(r,\phi,z) = \sum_n p_n(r,z)\cos n\phi \tag{22}$$

Because the cosines are orthogonal, each term in the series can be solved for separately. The pressure is governed by the Helmholtz equation in cylindrical coordinates, augmented by the non-axisymmetric differential operator $r^{-2}\partial^2/\partial\phi^2 = -r^{-2}n^2$. We proceed with the solution in terms of the Fourier transform of the acceleration extended only over the length 2L of the actual cylindrical source since $\ddot{w}=0$ over the semi-infinite baffles:

$$\tilde{\ddot{w}}_n(\gamma) = \int_{-L}^{L} \ddot{w}_n(z)e^{-i\gamma z}dz$$

The transform of the boundary condition, Eq. 5.16, takes the form

$$\left.\frac{\partial \tilde{p}_n(r,\gamma)}{\partial r}\right|_{r=a} = -\rho\tilde{\ddot{w}}_n(\gamma) \tag{23}$$

2. The Far-Field

The pressure transforms are governed by the Fourier-transformed Helmholtz equation

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{n^2}{r^2} + k^2 - \gamma^2\right)\tilde{p}_n(r,\gamma) = 0 \tag{24}$$

The solution of this equation are Bessel and Neumann functions of argument $(k^2-\gamma^2)^{1/2}r$. When the radial coordinate is large, the appropriate linear combination of these functions must display the phase angle of a plane wave. As seen from Eq. 12, this condition is satisfied by the Hankel function of the first kind, where the superscript will be dropped for the sake of conciseness. The desired general solution of Eq. 24 is therefore

$$\tilde{p}_n(r,\gamma) = A_n H_n[(k^2-\gamma^2)^{1/2}r] \tag{25b}$$

$$\simeq \frac{2^{1/2}A_n(-i)^n\exp]i(k^2-\gamma^2)^{1/2}r-i(\pi/4)]}{[\pi(k^2-\gamma^2)^{1/2}r]^{1/2}}, \quad (k^2-\gamma^2)^{1/2}r >> n^2+1 \tag{25a}$$

When Eq. 25a is substituted in the boundary condition, Eq. 23, one can solve for A_n in terms of $\tilde{\ddot{w}}_n$. The far-field pressure is finally constructed as an inverse Fourier transform of Eq. 25b. Previously, however, we switch to spherical coordinates $r = R\sin\theta$, $z = R\cos\theta$. The inverse Fourier transform now becomes

$$p(R,\theta,\phi) = \frac{-\rho}{(2\pi^3 R\sin\theta)^{1/2}}\sum (-i)^n\exp\left(-\frac{i\pi}{4}\right)\cos n\phi$$
$$\int_{-\infty}^{\infty} \frac{\tilde{\ddot{w}}_n(\gamma)\exp\{iR[(k^2-\gamma^2)^{1/2}\sin\theta+\gamma\cos\theta]\}}{(k^2-\gamma^2)^{3/4}H_n'[(k^2-\gamma^2)^{1/2}a]} \tag{26}$$

For large R, this integral, being of the form of Eq. 15 can again be evaluated by the method of stationary-phase, where $\bar{\gamma} = k\cos\theta$, $(k^2-\bar{\gamma}^2)^{1/2} = k\sin\theta$. Eq. 16 finally yields

$$p(R,\theta,\phi) = \frac{\rho e^{ikR}}{\pi kR\sin\theta}\sum_{n=0}^{\Sigma} \frac{\tilde{\ddot{w}}_n(k\cos\theta)(-i)^{n-1}\cos n\phi}{H_n'(ka\sin\theta)} \tag{27}$$

3. The Uniformly Pulsating Cylinder

Consider a source undergoing uniform radial vibrations. Only the n=0 term in Eq. 27 need be retained. The Fourier transform of the dynamic configuration is

$$\tilde{\ddot{w}}_o(\gamma) = \ddot{W}_o\int_{-L}^{L} e^{-i\gamma z}\,d\gamma = \frac{2\sin\gamma L}{\gamma} \tag{28}$$

When this is substituted in Eq. 27, with $\gamma = \bar{\gamma}$, one obtains the far-field

$$p(R,\theta) = \frac{i2\rho e^{ikR}\ddot{W}_o\sin(kL\cos\theta)}{\pi k^2 R\cos\theta\sin\theta H_o'[ka\sin\theta]} \tag{29}$$

In the low-frequency limit,

$$H_o' \simeq \frac{2i}{\pi ka \sin\theta} \quad , \qquad k^2a^2 << 1$$

Eq. 29 reduces to

$$p(R,\theta) \simeq \frac{\rho \ddot{W}_o a e^{ikR} \sin(kL\cos\theta)}{kR\cos\theta} \tag{30}$$

This coincides with the result obtained by elementary considerations for the line source, Eq. 3, with $k_o = 0$, $\ddot{Q} = 4\pi La\ddot{W}_o$, and p_s as given in Eq. 5.20.

4. Convex Sources in General

The cylindrical source is obviously periodic in ϕ, and therefore requires only a single Fourier transform. The sphere and the spheroid are doubly periodic in their two angular coordinates and therefore do not require any transforms. The same applies to the infinite cylinder whose dynamic configuration is periodic in both z and ϕ. Since the sound field can, in these cases, be constructed entirely in terms of wave-harmonic series without recourse to transforms, the difficulty of evaluating the inverse transform is circumvented. These configurations are therefore tractable in the near- as well as far-field. Unfortunately, convex bodies for which the wave equation is not separable do not allow either a wave-harmonic or a transform solution, but require a computer formulation except in the low-frequency limit.

REFERENCES

1 See for example C. J. Tranter, Integral Transforms of Mathematical Physics, 3rd ed. (New York, Wiley, 1966), p. 16.

2 Ibid, pp. 47-48.

3 H. Lamb, Hydrodynamics, 6th ed. (New York, Dover Publishing Co., 1945), pp. 395-396.

STOCHASTIC EXCITATION OF ELASTIC STRUCTURES AND EXAMPLES OF FLOW-GENERATED NOISE

William K. Blake
Research Scientist
DAVID TAYLOR NAVAL SHIP
R&D CENTER
Bethesda, Maryland 20084
UNITED STATES

Summary

In this lecture we will use the fundamentals of flow-generated sound, developed in Reference 1, to determine the important parameters that must be measured to make quantitative evaluations. Using as an example the flow-excited vibration of cylinders the principles will be related to actual measurements and, as well, the fundamentals of self-sustained vibrations will be introduced. The lecture will end with a brief survey of the important semi-empirical relationships that govern the more complex generation mechanisms of different types.

1.0 Introduction

In the lecture on Fundamental Concepts of Flow-Generated Noise[1], the general equations of the fundamental flow-induced noise sources were presented and these were used to establish basic rules of similitude. In this lecture we will examine, through the simple example of the Aeolian tone, the types of parameters which must be quantified in order to either predict noise levels or to establish the most sensitive variables through which noise control may best be effected. The second part of the lecture will review the basic characteristics of some important practical noise source types.

2.0 Sound Radiated By Flow Past A Flexible Cylinder

2.1 Sound From Flow Dipoles

We assume that the cylinder is positioned in free space; i.e. away from any acoustically reflecting boundaries as shown in Figure 1. Its diameter is assumed to be much smaller than an acoustic wave length and the axial correlation length $2\Lambda_3$ of the forces on the cylinder is much shorter than the length of the cylinder L exposed to the flow. The unsteady lift forces $F_2(t)$ are nearly tonal and generated at the frequency f_s that corresponds to the frequency of shedding of the vortices. Since the fluctuating drag forces are roughly only 1/10 the fluctuating lift we shall ignore the drag forces and consider only flow-induced sound and vibration in response to lift fluctuations. Although the forces are temporally tonal, they are random functions of positions along the length of the cylinder. Accordingly, the cross spectral density of the forces per unit length at points y_3 and y_3' may be represented as [2,3]

$$\Phi_F(y_3', y_3, \omega) = \frac{1}{2} q^2 d^2 \overline{C_L^2}\, R(y_3, y_3)\, \delta(\omega \pm \omega_s) \tag{1}$$

So that the mean square lift per unit length is

$$\overline{f_2'^2} = q^2 d^2 \overline{C_L^2} \tag{2}$$

The correlation function $R(y_3', y_3)$ expresses the spatial statistics of the lift per unit length along the cylinder; $\overline{C_L^2}$ is a coefficient of mean-square lift, $q = \rho_0 U^2/2$ is the dynamic pressure and d is the diameter of the cylinder. Figure 2 shows an illustration of the correlation function which is assumed to be independent of any end points; i.e. we assume that $R(y_3', y_3) = R(y_3' - y_3) = R(r_3)$ and $R(0) = 1$ within the region of flow. The correlation function illustrated applies to rigid cylinders only. In the next section we shall examine conditions under which the forces may be influenced by the vibration of the cylinder. Although these influences may be considerable the relationship between the dipole and vibration contributions should be unchanged. The sound is a superposition of two contributions: the direct dipole sound due to the forces $F_2(t)$ and the sound due to the flow-induced vibration. The point of this exercise to examine the conditions for which one contribution dominates over the other.

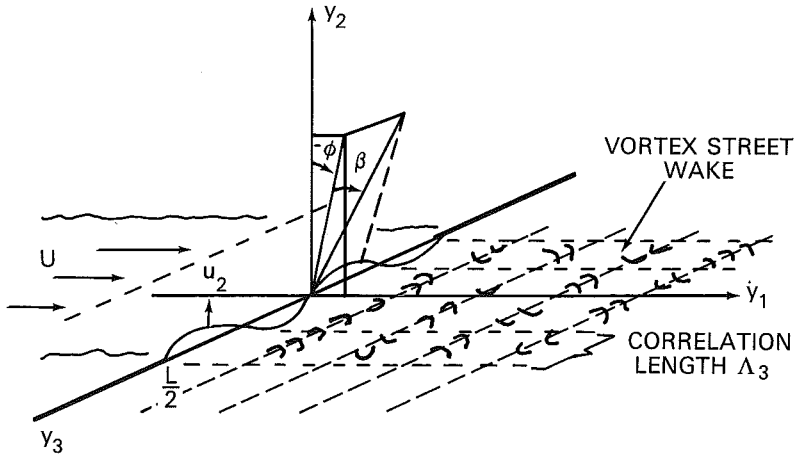


Figure 1 - Idealization of Flow-Induced Vibration of a Thin Flexible Segment of a Long Cylinder of Greater Length. Flow Extends Along the Flexible Portion Only.

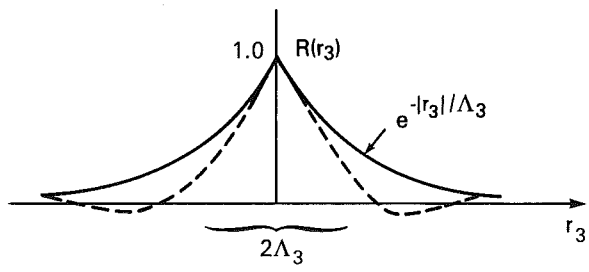


Figure 2 - Typical Correlation Functions for Flow-Induced Lift on Rigid Cylinders.

Equation (9) of "Fundamental Concepts of Flow Generated Noise", ref. 1, yields an expression for the direct dipole sound

$$\left[\Phi_{p_r}(r,\phi,\omega)\right]_{DIPOLE} = \frac{(k_o d)^2 \rho^2 \cos^2\phi}{16\pi^2 r^2} \, 2\Lambda_3 L \, \overline{C_L^2}\left\{\frac{1}{2}\delta(\omega \pm \omega_s)\right\} \qquad (3)$$

where $k_o = \omega/c_o$ and $k_o \Lambda_3 << 1$. This equation is easily cast in terms of equations 16 of ref. 1 by noting that the mean square sound pressure (eq. 6, ref. 1) is dominated by frequency $\omega = \omega_s = 2\pi S U/d$ where $S \simeq 0.21$ is the Strouhal number of the vortex shedding. The function D_o is, now, the factor $10\log[\overline{C_L^2}\cos^2\phi/16\pi^2]$. A_c/L^2 is replaced by $2\Lambda_3/L$. Note that the radiated sound pressure is fully quantified by measurements of $\overline{C_L^2}$, S and Λ_3 . If vibration of the cylinder influences the forces the relationship (3) still applies, but appropriate values of $\overline{C_L^2}$ and Λ_3 must be known. Some typical values of these parameters as a function of Reynolds number for rigid cylinders are shown in Figure 3.

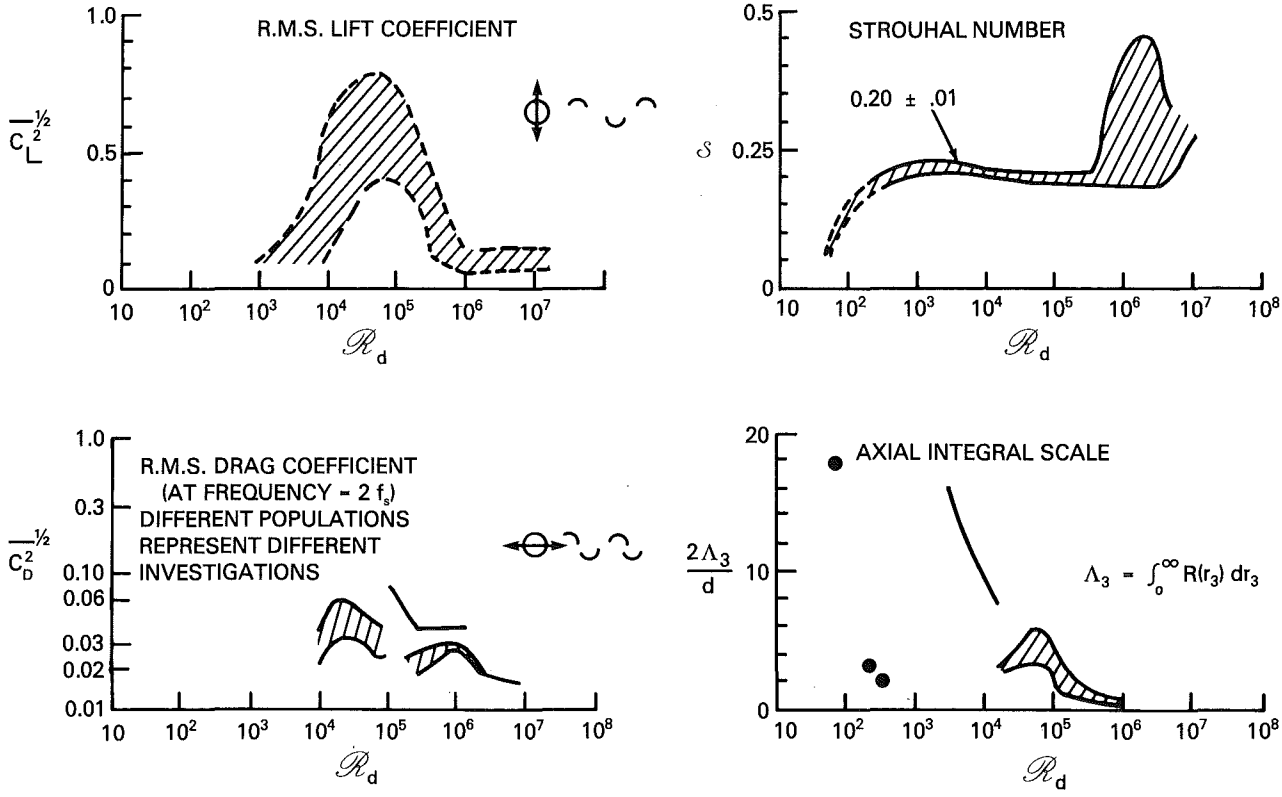


Figure 3 - Summary of Parameters for Vortex-Induced Forces on Circular Cylinders.

2.2 Sound From Flow-Induced Vibration

The flow-induced vibration can be determined following Section 5.0 in ref. 1. Methods for doing this may be found in the literature[25]. Thus, equation (28) of ref. 1 gives the cross-spectral density of the velocity u_2 of the string vibrating in its n th mode

$$\left[\Phi_{u_2}(y_3', y_3, \omega)\right]_n = \frac{\omega^2 \psi_n(y_3)\psi_n(y_3')}{M^2 \omega_n^4 |\mathfrak{Z}|^2} \rho^2 d^2 \overline{C_L^2} \left[\frac{1}{2}\delta(\omega \pm \omega_s)\right] \cdot \int_{-L/2}^{L/2}\int_{-L/2}^{L/2} R(r_3)\psi_n(z_3)\psi_n(z_3 + r_3)\, dr_3\, dz_3 \qquad (4)$$

and $$\left[\Phi_{u_2}(y_3', y_3, \omega)\right] = \sum_n \left[\Phi_{u_2}(y_3', y_3, \omega)\right]_n$$

the function $|\mathfrak{Z}|$ is given by equation (29) of ref. 1. Utilization of the Fourier transform permits an alternative representation.

$$\left[\Phi_{u_2}(y_3', y_3, \omega)\right]_n = \frac{\psi_n(y_3')\,\psi_n(y_3)\,\omega^2}{M^2\omega_n^4\,|3|^2}\,\rho^2 L^2 d^2\,\overline{C_L^2}\left[\tfrac{1}{2}\delta(\omega \pm \omega_s)\right]\int_{-\infty}^{\infty}\phi(k_3)\,|S_n(k_3)|^2\,dk_3 \tag{5}$$

where the modal shape function is

$$S_n(k_3) = \int_{-\infty}^{\infty} e^{ik_3 y_3}\,\psi_n(y_3)\,dy_3$$

and the wave number spectrum is

$$\phi(k_3) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ik_3 r_3}\,R(r_3)\,dr_3$$

The key to the equivalence of equations (4) and (5) is the presumption of spatial stationary of the lift forces along the cylinder, i.e. that $R(r_3)$ is independent of axial location of the origin y3. Without this assumption it may be more fruitful to work with equation (4).

To determine the sound pressure at distances which are far from the cylinder we must evaluate the integral over y_s in equation (35) of ref 1. Since the source-to-field distance may be approximated by (see Figure 1)

$$|\vec{x} - y_3| \simeq r - \frac{d}{2}\cos\beta\,\cos(\phi - \phi_0) - y_3\sin\beta$$

and

$$u_n = u_2\cos\phi_0$$

the field acoustic pressure is determined by

$$\left[\Phi_{p_{rad}}(r, \beta, \phi, \omega)\right]_n = \frac{k_0^2\cos^2\beta\,\cos^2\phi}{16\pi^2 r^2}\left[\frac{m_a}{m_c + m_a}\right]\frac{\rho^2 d^2 C_L^2}{|3|^2 L^2}\left[\tfrac{1}{2}\delta(\omega - \omega_s)\right]\cdot$$
$$|S_n(k_0\sin\beta)|^2\int_{-\infty}^{\infty}\phi(k_3)\,|S_n(k_3)|^2\,dk_3 \tag{6}$$

where the $|S_n(k_3)|$ is given above, m_a is the added mass per unit length, $\rho_0\pi d^2/4$, and m_c is the cylinder mass per unit length, $\rho_c\pi d^2/4$.

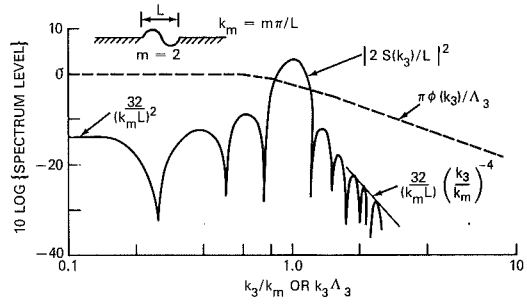


Figure 4 - Modal Shape Function for a Simply-Supported Structure and Wave-Number Spectrum of Unsteady Lift Per Unit Length.

Both the vibration response and the far field sound pressure depend on the spatial matching of the lift fluctuation and the mode shape. In equations (5) and (6) this is expressed in terms of the wave number spectra $\phi(k_3)$ and $S_n(k_3)$. These are illustrated in Figure 4; $S(k_3)$ is peaked about $k_3 = k_n = n\pi/L$ and $\phi(k_3)$ is bounded by $k_3 < 1/\Lambda_3$. When $\Lambda_3 << 2\pi/k_n = 2L/n$ then the integral may be approximately by $2\Lambda_3 L$. Accordingly for correlation lengths which are much smaller than a wavelength λ_n of vibration we have for each of the n modes

$$\left[\Phi_{p_{rad}}(r, \phi, \beta, \omega)\right]_n = \frac{(k_0 d)^2\,\rho^2\cos^2\phi\,\cos^2\beta}{16\pi^2 r^2}\,2\Lambda_3\,\overline{C_L^2}\left[\tfrac{1}{2}\delta(\omega \pm \omega_s)\right]\cdot$$
$$\left\{\left(\frac{\rho_0}{\rho_0 + \rho_c}\right)^2\frac{1}{|3|^2}\,\frac{1}{L^2}\,|S_n(k_0\sin\beta)|^2\right\}. \tag{7}$$

2.3 Relative Importance of Dipole and Vibration Components

The term in curley brackets represents the structural acoustic coupling of the cylinder motion with the acoustic medium; this coupling includes the structural admittance of the cylinder as indicated by the presence of $|\mathfrak{Z}|$ in the denominator. Comparison with equation (3) also shows that this bracketed term also governs the relative importance of the vibration-induced sound and the direct dipole sound. According to Figure 4, if $k_o << k_n$ the $|S_n(k_o \sin\beta)|^2$ can be approximated by its asymptotic value as $k_o \to 0$ thus

$$|S_n(k_3 \to 0)|^2 \sim \beta/k_n^2 \tag{8}$$

so that the ratio of the vibration-induced sound to be dipole sound is

$$\frac{\left[\Phi_{p_{rad_n}}(r,\phi,\beta,\omega)\right]_{VIB}}{\left[\Phi_{p_{rad}}(r,\phi,\omega)\right]_{DIPOLE}} \simeq \frac{\cos^2\beta}{|\mathfrak{Z}|^2}\left(\frac{\rho_o}{\rho_o+\rho_c}\right)^2 \frac{1}{L^2}|S_n(k_o \sin\beta)|^2 \tag{9}$$

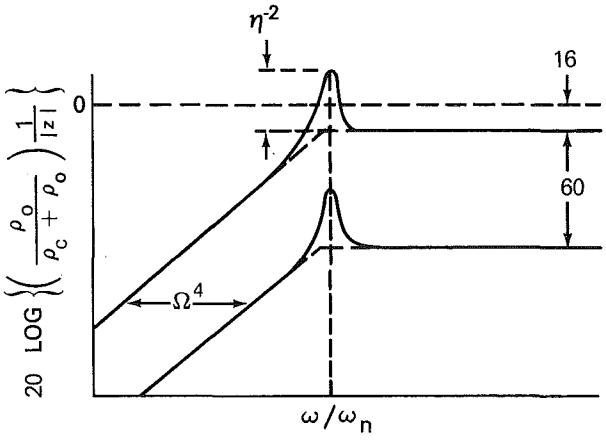


Figure 5 - Ratio of $[\Phi]_{VIB}/[\Phi]_{DIPOLE}$ for ¼ Inch Dia. Steel Cylinders in Air and in Water.

Figure 5 shows this ratio as a function of ω/ω_n for a steel cylinder in air and in water. It shows that in air, the vibration-induced contribution is not dominant, while it is in water. The evaluation is made to give an upper bound for which $k_n=k_o$ and $\cos\beta = 1$, so $|S_n(k_o \sin\beta)|^2 = 2$.

2.4 General Characteristics of Structural Response and Noise Control

This simple example illustrates many important features of flow-induced noise. First, in aeroacoustic problems the vibration-induced contribution is often minimal importance compared with the direct dipole contribution, while the reverse is true in liquids. This is because the impedance coupling (in this case represented by $[\rho_o/(\rho_o+\rho_c)]^2$) is quite small. Second, the dipole sound will be controlled by wave numbers that are less than or equal to the acoustic wave number. In equation (3), this would be stated by replacing $2\Lambda_3$ by $2\pi\,\phi(k_3 = k_o \sin\beta)$. This replacement actually gives the more general representation that applies to all manner of flow source that is governed by equations (8) and (10). Specifically, the acoustic radiation is related to the gross features of the spatial correlation function of the sources, or rather to the very low wave number spectrum of the source region. Similarly, the structural response, and the associated acoustic radiation are governed by the wave number spectrum of the source at wave numbers near those of waves in the structure for the vibration level and wave numbers near the acoustic for the sound level.

The implication to measurement programs is that details of the flow which account for the obvious correlation is the flow field (of turbulence, say) may bear little relationship to the problem of flow excitation.

Consider, for example the response of two dimensional panels such as in aircraft fuselages to turbulent boundary layers. For frequencies greater than the hydrodynamic coincidence frequency, $\omega > \omega_h = U_c^2 / C_L h / \sqrt{12}$ (i.e. the frequency for which free bending waves in the structure equals the pressure convection velocity) the panel response is controlled by a two dimensional integral just like that in equation (5) giving the a general form for the mean-square surface velocity of frequency bands that include multiple resonances of the panel

$$\overline{V^2} = \frac{\pi}{m_s^2 \eta \omega} \left\{ \int_{-\infty}^{\infty}\int \Phi_p(\vec{k},\omega) \left| S_{mn}(\vec{k}) \right|^2 d^2\vec{k} \right\} n(\omega) \Delta\omega \qquad (10)$$

where $\Phi_p(\vec{k},\omega)$ is strongly peaked at $k_3=0$ and $k=k_c=\omega/U_c$ and $|S_{mn}(\vec{k})|^2$ is peaked near $k_{mn}=k_p$. Since $k_c >> k_p$ for this case, the part of $\Phi_p(\vec{k},\omega)$ that couples with $|S_{mn}(\vec{k})|^2$ is that which has relatively small values at the low wave number; m_s is the plate mass per unit area. The $n(\omega)$ is the mode density and for flat plates it is $n(\omega) = \sqrt{12}\, A_p / \pi h C_L$ where h is the thickness, A_p is the area, and C_L is the longitudinal wave speed in the plate.

Measurement programs which are directed at providing information for the outright prediction of flow-excited vibration and noise must therefore be conducted with these notions firmly in mind. It is necessary to conduct the flow measurements in a manner that is consistant with the mechanism of flow-structure or flow-acoustic coupling. In a very real sense, in the flow-induced noise generation energy is transducted from the aerodynamics into acoustic emissions. Both the structure and the acoustic medium act as filters in this process. From this perspective noise control may be effected by any of these means:

- Reduce the magnitude of $\Phi_p(\vec{k},\omega)$ or the distribution of $\Phi_p(\vec{k},\omega)$ over $\vec{k}$ by modifying the flow.
- Reduce the response of the structure for a given flow input by altering $|S_{mn}(\vec{k})|$ by increasing its mass M, or by increasing the damping η
- Reduce the radiation efficiency by altering $S_{mn}(k_o \sin\beta)$ which, according to equation (35) of reference 1, is related to σ_{rad} by

$$\sigma_{rad} \sim \int_{\beta=0}^{\pi} \int_{\phi=0}^{2\pi} (k_o d)^2 \cos^2\phi \cos^2\beta \left| S_n(k_o \sin\beta) \right|^2 \sin\phi \, d\beta \, d\phi \qquad (11)$$

It is to be noted that considerations of noise control by damping must be tempered by recognition of the various damping mechanisms.

In the simple example, the damping of the cylinder would be controlled by the mechanical losses η_m (probably at the end points of the cylinder) and by radiation damping η_{rad} (because acoustic energy is radiated away from the cylinder). Therefore we may write

$$\eta = \eta_m + \eta_{rad} \qquad (12a)$$

where

$$\eta_{rad} = \frac{4 \sigma_{rad} \rho_o C_o}{\rho_p d \omega} \qquad (12b)$$

for the cylinder. Accordingly additions of hysteretic damping will only be effective when $\eta_m > \eta_{rad}$.

These notions generally hold true for more complex structural geometries.

2.5 Self-Excited Vibrations of the Cylinder

The cylinder also provides an example of self-sustained oscillations and the parameters which control them. For values of cylinder displacement (in the lift direction) that are greater than approximately (0.01)d peak to peak, or $\overline{\xi_2^2}^{1/2} > (0.007)d$ one can usually expect some nonlinear interaction between the flow forces and the vibration. This is especially true when the frequency of vortex shedding is equal to the frequency of vibration. Such nonlinearities and self-excited vibrations occur when $\omega_s \simeq \omega_n$. Figure 6 shows an example of a lift coefficient obtained experimentally as the cylinder

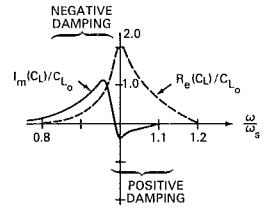


Figure 6 - Behavior of Lift Coefficient, $C_L = \frac{1}{qL}\int_o^L f_2\,(y_3)\,dy_3$, as Function of Frequency of Cylinder Vibration at a Displacement Amplitude Equal to 0.01 d. C_{L_o} is RMS Value on Stationary Cylinder at $\omega = \omega_{Si}\mathscr{R}_d = 3$ to $10X10^6$.

was mechanically oscillated in the cross-stream direction at constant amplitude and at varying frequencies above and below the natural frequency of vortex shedding. When the vibration couples with the wake, the lift coefficient induced by the vortices becomes phase-locked with the motion of the cylinder. For the convention used, the positive imaginary lift coefficients are taken by the structure as negative damping (energy fed from the wake into the cylinder). Also the real part (in phase with the displacement) of the lift increases to a value that is greater than that for the rigid cylinder. The flow-induced vibration of the cylinder will therefore greatly depend on the relative values of $\eta m_c \omega_n^2 \xi_2$ and $-q\,l_m(C_L)$. Assuming a relationship between $\frac{I_m(C_L)}{\xi_2/d}$ and $R_e(C_L)$, and displacement amplitude divided by cylinder diameter, then for a given cylinder geometry and flow Reynolds number, a relationship[4,5]

$$\frac{\xi_2}{d} = f\left(\frac{\rho_o d^2}{\eta_m m_c}\right) \tag{13}$$

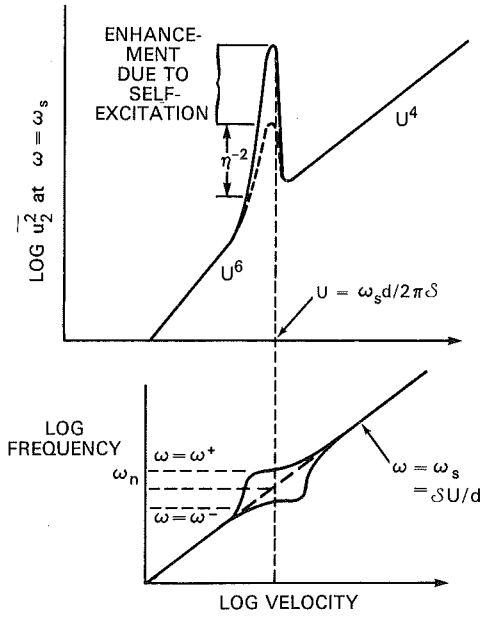


Figure 7 - Flow-Induced Vibration at Vortex Shedding Frequency and Frequency of Vortex Shedding as a Function of Velocity for an Elastic Cylinder or Hydrofoil in Cross Flow.

can be found. Such relationships have been found for circular cylinders and for hydrofoils with squared-off trailing edges. An illustration of type of self-excited motion due to vortex shedding to be expected is shown in Figure 7. In the vicinity of $\omega_s \simeq \omega_n$ the frequency of vortex shedding and of cylinder motion become coupled so that $\omega = \omega^+$ or $\omega = \omega^-$ become the resonance frequencies of the coupled flow - cylinder system. The attainment of ω^+ or ω^- depends on the degree of flow-structure coupling as controlled by the mechanical damping.

Figure 8 is an illustration of the self-sustained amplitude of singing hydrofoils with blunt trailing edges.

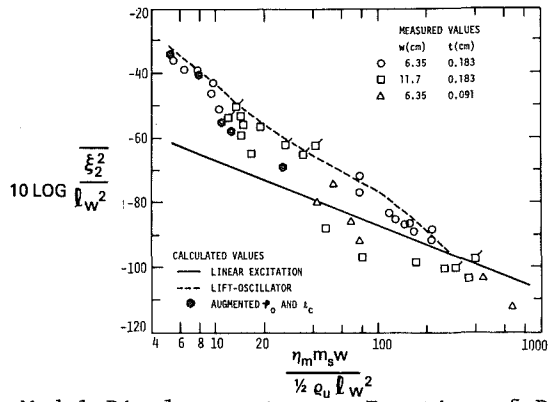


Figure 8 - Modal Displacement as a Function of Damping Parameter and ℓ_w^2 for Modes of Hydrofoils with Various Trailing Edge Thicknesses and Chords.

One would expect similar, though possibly more or less pronounced, behavior for hydrofoils with edges of different geometries. In the case of vibration of hydrofoils the length scale l_w replaces d for the cylinder. l_w refers to the thickness of the shear layer in the wakes of the hydrofoil near its trailing edge and it controls the vortex shedding frequency such that[5]

$$\frac{\omega_s \ell_w}{U\sqrt{1-(C_p)_b}}$$

where $(C_p)_b$ is the base pressure coefficient.

3.0 Effects of Acoustically Non-Compact Surfaces On Dipole Sound

Equation (3) gives the sound pressure due to the total mean-square force of $\overline{C_L^2}\, q^2 d^2 (2\Lambda_3 L)$ in which the diameter d and the $2\Lambda_3$ are to be much smaller than an acoustic wave length. The first note of generalization applies to a longer correlation length compared to an acoustic wave length. This generalization was noted earlier in Section 2.4 by the replacement of $2\Lambda_3$ by $2\pi\, \phi(k_3 = k_o \sin\beta)$

Since the axial correlation functions of vortex-induced forces, such as measured on rigid cylinders, are representable as $\exp[-|\zeta_3|/\Lambda_3]$ the correction to the dipole sound due to non-compactness of the correlation is

$$\frac{\pi}{\Lambda_3}\, \phi_3(k_3 = k_o \sin\beta) = \frac{1}{1 + (k_o \Lambda_3 \sin\beta)^2} \tag{14}$$

Another level of non-compactness applies when the stream-wise extent of the body is longer than an acoustic wave length, as is often the case with lifting surfaces. In such cases, a new acoustic Green function for a rigid half plane to be used in equation (12) Ref. 1 gives rise to an alternative formula

$$\Phi_{p_{rad}}(\vec{x},\omega) \sim (k_o \ell_o)\,|\sin\beta|\,\sin^2\phi/2\;\frac{2\Lambda_3 L}{r^2}\, q^2\, \phi_p(\omega) \tag{15}$$

where l_o is a measure of the location of the dipole source region relative to the edge and $\phi_p(\omega)$ is a spectrum function (with dimensionaless integral over frequency) of the sources. Equation 15, gives a $U^4 (U/c_o)$ dependence of overall acoustic intensity compared with a $U^4 (U/c_o)$ dependence given by equation 3. Accordingly, overall sound pressures from direct dipole sound from rigid airfoils will increase as U^5 whenever the ratio chord$/\lambda_o$ is greater than unity. The effective distance of sources from the edge l_o is related to the dimension of the source region, say the thickness of the boundary layer of flow near the trailing edge δ. Therefore, the parametric expression of sound pressure spectrum will replace $k_o l_o$ with

$$k_o l_o \sim \left(\frac{\omega \delta}{U}\right)\left(\frac{U}{c_o}\right)$$

4.0 Survey of Characteristics of Important Flow-Induced Noises

Using the above fundamentals, the essential behavior of many practical noises can now be put into perspective. The accompanying table summarizes a number of types of sources which are classified according to: A, monopole; B, dipole, and C, quadrupole. A forth category, D, covers free shear-layer tones due to flow impingement on edges and covers. Where data is available, an expression for radiated sound power is given. In the cases for monopole, dipole, and quadrupole sounds, the expressions for radiated sound power are all consistent with the expressions derived in this lecture.

In the cases of flow tones, all sources are dipole like, but little is known about the acoustic amplitudes because the pressure amplitudes depend on the details of each feed back mechanism that sustains the tone. The feed back is generally controlled by an acoustic path between the upstream and downstream edges. In hydroacoustic applications, however, the feed back is generally controlled by a structural vibration path.

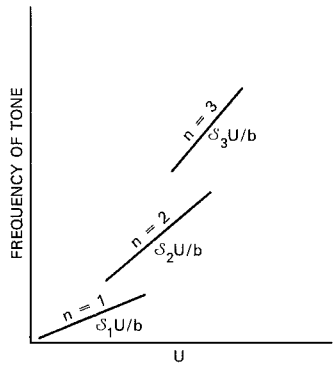


Figure 9 - Illustration of the Frequency-Speed Relationship for Staged Tones.

The behavior of staged tones; i.e. those with n=1,2,3... etc. modes, have the frequency-speed relationship has jumps as illustrated in Figure 9. The end points by each stage usually depend on whether the velocity increases or decreases. Two or three stages are generally typical. No universal criteria, or explanations, for that manner, have been offered to account for the speed range for which any particular mode exists. A mode is generally regarded as being established by the number of quarter waves of free shearlayer modes which "fit" between the leading and trailing edges of the aperture consistently with certain boundary conditions. Both viscous and acoustic effects must control the existance or non-existance of these modes. Howe has recently attributed the speed range of resonance of a Helmholtz resonator reinforcing particular modes of the free shear-layer in the opening of the resonator as being controlled by acoustic effects.

5.0 References

1. Blake, W.K. "Fundamental Concepts of Flow-Generated Noise" Lecture Notes

2. Phillips, O.M. "The Intensity of Aeolian Tones" J. Fluid Mech. 1, 607-624, 1956

3. Leekey, P. and Hanson, C.E. "Aeolian Tones Associated with Resonant Vibration" J. Sound Vibration. 13, 465-483, 1971

4. Skop. R.D. and Griffin, O.M. "On A Theory for the Vortex-Excited Oscillations of Flexible Cylindrical Structures" J. Sound Vib. 41, 263-274, 1975

5. Blake, W.K., Maga, L.J., Finkelstein, G. "Hydroelastic Variables Influencing Propeller and Hydrofoil Singing" ASME Symposium on Noise and Fluid Engineering, 1977

6. Ross, D. "Mechanics of Underwater Noise" Permagon Press, 1976

7. Blake, W.K. Wolpert, M.S., Geib, F.E. "Cavitation Noise and Inception as Influenced By Boundary Layer Development on a Hydrofoil" J. Fluid Mech. 80, 617-640, 1977

8. Hubert, M. "Stromungsgerauche" Verein Deutscher Ingenieure VDI - Bilddenginswerk, 1-21, 1969

9. Gongwer, C.A. "A Study of Vanes Singing In Water" Trans ASME, J. Appl. Mech 19, 432-438, 1952

10. Chi, M. "Response of Bridge Structural Members Under Wind Induced Vibrators" U.S. Federal Highway Administration Report. FHWA-RD-78-25, 1978

11. Weaver, D.S. and El-Kashlan, M. "On the Number of Tube Rows Required To Study Cross-Flow Induced Vibrations in Tube Banks" J. Sound Vib. 75, 265-273, 1981

12. Blevens, R. "Flow-Induced Vibrations" Van Nostrand - Reinhold, 1977

13. Graham, J.M.R. "Lifting Surface Theory For The Problem of An Arbitrarily Yawed Sinusoidal Gust Incident on a Thin Airfoil in Incompressible Flow" Aero. Quarterly, 21, 182-198, 1969

14. Mugridge, B.D. "The Noise of Cooling Fans Used In Heavy Automotive Vehicles" J. Sound Vib. 44, 349-367, 1976

15. Howe, M.S. "A Review of the Theory of Trailing Edge Noise" J. Sound Vib. 61, 437-465, 1978

16. Tam, C.K.W. Reddy, N.N. "Sound Generated In the Vacinity of The Trailing edge of an Upper Surface Blown Flap" J. Sound Vib. 52, 211-232, 1977

17. Fink, M.R. "A Method For Calculating Externally Blown Flap Noise" NASA CR2954, 1978

18. Wilby, J.F., Gloyna, F.L. "Vibration Measurements of an Airplane Fuselage Structure I. Turbulent Boundary Layer Excitation" J. Sound Vib. 23, 443-466, 1972

19. Goldstein, M.E. "Aeroacoustics" McGraw Hill, 1976

20. Reethoff, G. "Control Valve and Regulator Noise Generation, Propagation, and Reduction" Noise Control Eng. 9, 74-85, 1977

21. Powell, A. "On the Edge Tone" J. Acoustical Soc. Am. 33, 395-409, 1961

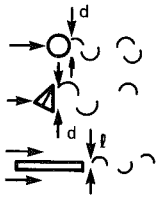
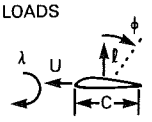
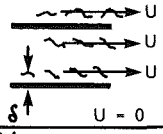
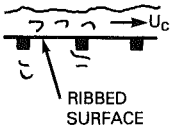
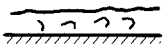
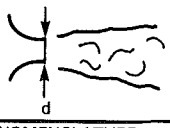
22. Chanaud, R. and Powell, A. "Some Experiments Concerning the Hole and Ring Tone" J. Acoust. Soc. Am. 37, 902-911, 1965

23. Rockwell, D. and Naudasher, E. "Self-Sustained Oscillations of Impinging Free Shear Layers" Ann. Rev. Fluid Mech 11, 67-94, 1979

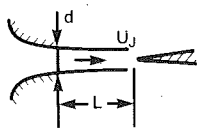
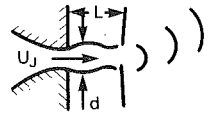
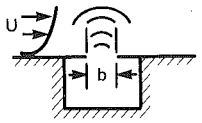
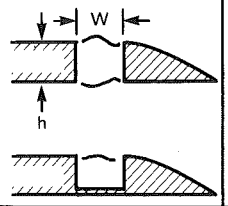
24. Jones, G.W. "Unsteady Forces Generated By Vortex Shedding About a Large Stationary and Oscillating Cylinder at High Reynolds Number" ASME paper 68-FE-36, 1968

25. Blake, W.K. "Aero-Hydroacoustics For Ships" 2 Vols., DTNSRDC Report 80-001, 1982 To Be Published

TABLE I IMPORTANT SOURCE TYPES AND THEIR PARAMETERS

	SOURCE TYPE	APPLICATION	CHARACTERISTIC OR FUNDAMENTAL FREQUENCY	VIBRATION & SOUND DESCRIPTORS
MONOPOLE	A) CAVITATION 	HYDRAULIC MACHINERY[6,7] VALVES THROTTLES	BROADBAND $f_c \sim \frac{1}{L}\sqrt{\frac{P_\infty}{\rho_o}}\,\frac{1}{\sqrt{Ki-K}}$	$\mathbb{P}_{rad}(w) \sim \frac{P_\infty^2 L^2 (Ki-K)^m}{\rho_o C_o}$, $m \geq 1$
DIPOLE	B-1 VORTEX SHEDDING 	PIER PILES[4] GRATINGS[8], PROPELLER SINGING[9], STRUCTURAL MEMBERS[10] TUBE BUNDLES,[11] BUILDING VIBRATIONS,[12] ANTENNAS	TONAL $\frac{fd}{U} = 0.22$ $\frac{fd}{U} \cong 0.2$ $f\ell/U \cong \frac{1}{2}\pi$	EQS. (3) TO (10)
	B-2 FLUCTUATING LOADS 	FANS/COMPRESSORS[12,14]	BROADBAND & TONAL $f_c \sim \frac{U}{\lambda}$	$\mathbb{P}_{rad} = \frac{k_o^2}{12\pi}\frac{\overline{\mathcal{L}^2}}{\rho_o C_o}$ for $k_o C < \pi$
	B-3 TURBULENT TRAILING-EDGE NOISE 	LIFTING SURFACES,[15] BLOWN FLAPS,[16,17] MACHINERY NOISES	BROADBAND	$\mathbb{P}_{DIP} \cong \frac{\pi^2}{2}\left(\frac{q^2}{\rho_o C_o}\right) M_c \lambda_3 \delta \Phi\left(\frac{fd}{U}\right)$ $[\mathbb{P}]_{FLEX} \cong [\mathbb{P}]_{DIP}/4\beta;\ \beta = \frac{\rho_o C_o}{m_s 2\pi f} > 1$ Figure 10 for $\Phi\left(\frac{fd}{U}\right)$
	B-4 FLOW—INDUCED VIB. 	AIRCRAFT CABIN NOISE,[18] AND VIBRATION	BROADBAND	$\overline{V^2} \cong \frac{4}{\pi}\frac{\Phi p(\omega)\Delta\omega}{m_s^2 \omega^2 \eta_t}\frac{\gamma_3}{(\gamma_3^2+1)}\left(\frac{\omega_h}{\omega}\right)^{1/2}$ $\gamma_3 \sim 0.7,\ \omega < \omega_h \cong U_c^2/\kappa C_\ell$ $\overline{V^2} \cong \frac{\pi^2}{m_s^2 \eta_t \omega}\frac{\Phi_p(k_p \ll \omega/Uc)}{\kappa C_\ell}$ $\omega > \omega_h$ $\mathbb{P}_{rad} = \rho_o \omega_o A_p \sigma_{rad} \overline{V^2}$
QUADRUPOLE	C-1 HOMOGENEOUS TBL ON FLAT SURFACE 	NOT OBSERVED	BROADBAND	$\mathbb{P}_{rad} \sim \frac{q^2}{\rho_o C_o} M_c^4 \delta^2$
	C-2 TURBULENT JETS 	VENTS, AIRCRAFT,[19] VALVES[20] CERTAIN MACHINERY $\frac{Ud}{\nu} > 10^4$	BROADBAND $\frac{fd}{U_j} \cong 0.4$	$(\mathbb{P}_{rad})_{OA} \cong 10^{-5}\frac{q^2}{\rho_o C_o} M_c^4 D^2$
	NOMENCLATURE	U = velocity U_c = addy connection velocity $\cong 0.6U$ d = diameter U_J = jet velocity P_∞ = static ambient pressure $M_c = U_c/c_o$	$\mathbb{P}_{rad}$ = sound power in proportional frequency band, $\Delta f \propto f$ $(\mathbb{P}_{rad})_{OA} = \frac{1}{2}\rho_o U^2$, dynamic pressure ν = kinematic viscosity of fluid	L = body length scale K = cavitation index $(P_\infty - P_v)/q$ P_v = liquid vapor pressure $\omega = 2\pi f$ $\overline{V^2}$ = Mean-square surface velocity

SOURCE TYPES, CONTINUED

SOURCE TYPE	APPLICATION	CHARACTERISTIC OR FUNDAMENTAL FREQUENCY	VIB'N & SOUND DESCRIPTORS
D-1 EDGE TONE 	LAMINAR RECTANGULAR JET IMPINGING ON SHARP EDGE[21], MUSICAL INSTRUMENTS	Tonal $\frac{f_s L}{U_j} \cong \frac{1}{2}(n + \frac{1}{4})$ $n = 1, 2 \ldots, 4 < L < 20$	$\left(\mathbb{P}_{rad}\right)_{OA} \simeq \frac{5}{3\pi}\left(\frac{\omega_s}{C_0}\right)^2 \frac{q^2}{\varrho_0 C_0} A_j^2$ $\omega_s = 2\pi f_s$
D-2 HOLE/RING TONE 	STAGED THROTTLING DEVICES[22]	Tonal $0.011\sqrt{\mathscr{R}_d} > \frac{fd}{U_j} > 0.5$ $\mathscr{R} < 3000$ $\frac{L}{d} = 2, 2.5, 3$	
D-3 CAVITY TONE 	CAVITIES[23], BLEED PORTS	$\frac{fb}{U} \cong 0.022\frac{b}{\theta}$ Lam b.l. $\frac{fb}{U} = 0.33(n - \frac{1}{4})$ $n = 1, 2 \ldots$ Turb. b.l.	$P_{rad} \sim P_{cav}\left(\frac{\omega A_{or}}{C_0 n}\right)$
D-4 GAPS 	FLAPS ON LIFTING SURFACES	$\frac{fW}{U} \sim$ (0.6 to 0.7)(n - ½) $n = 1, 2 \ldots$ $\frac{fW}{U} \sim (0.5 - 0.6)(n - \frac{1}{4})$ $n = 1, 2 \ldots$	

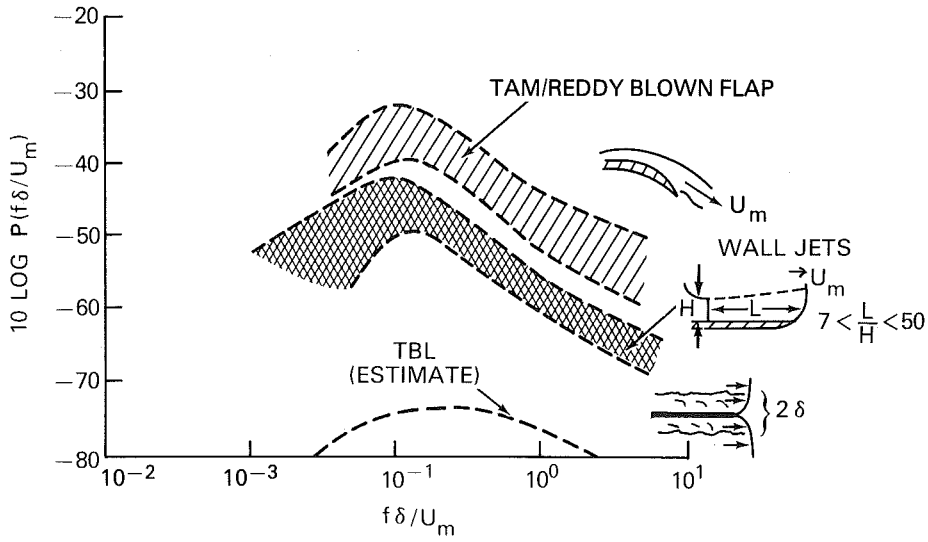


Figure 10 - One-Third Octave Band Levels for Trailing Edge Noise, Various Configurations
= Shear Layer Thickness at Edge.

SOUND RADIATION BY ELASTIC STRUCTURES

M. C. Junger
Cambridge Acoustical Associates, Inc.
54 Rindge Avenue Extension
Cambridge, Massachusetts 02140, USA

This lecture deals with realistic finite-impedance sound radiators whose vibrations are modified by the ambient fluid's radiation loading. Source-fluid interaction is illustrated for the elementary case of a pulsating gas bubble. The balance of the lecture deals with sound radiation by flexural waves in plates. The flexural wave velocity in vacuo and in the radiation-loaded plate is derived. Finally, a full-fledged interaction problem is solved by the transform technique introduced in Lecture 7: The sound field radiated by a force-excited effectively infinite elastic plate is constructed, radiation loading being rigorously accounted for.

So far we have implicitly assumed that the radiating surface has an infinite internal impedance in that its motion is not altered by the pressure exerted by the ambient medium. Actually, of course, this is not the case. One must therefore solve the elastic and acoustic problems simultaneously. We shall start with an almost trivial example.

A. AN ELEMENTARY INTERACTION PROBLEM

Our first example is extremely simple because the elastic system is merely a spring, viz., a spherical bubble in a liquid. Its spring stiffness is

$$K = -4\pi a^2 p/w \tag{1}$$

Introducing the bulk modulus, Eqs. 5.3 and 5.5, and noting that $dV/V = 3w/a$, the stiffness becomes

$$K = 12\pi a\gamma P_s \tag{2}$$

This is the solution of the elastic problem. The acoustic problem is solved in Eq. 5.30, which gives the required specific acoustic impedance components. Formulating them as mechanical impedances, and expressing the reactance in terms of a mass as $-i\omega M$, the small-ka asymptotic forms of the resistance and mass are

$$R = 4\pi k^2 a^4 \rho c$$
$$M = 4\pi a^3 \rho \tag{3}$$

The natural frequency of the bubble is therefore

$$f_o = \frac{1}{2\pi}\left(\frac{K}{M}\right)^{1/2} = \frac{1}{2\pi a}\left(\frac{3\gamma P_s}{\rho}\right)^{1/2} \tag{4}$$

For atmospheric pressure, f_o is of order 650 Hz over bubble diameter in cm. The resonance quality factor associated with radiation damping only is

$$Q = \frac{(KM)^{1/2}}{R} = \left.\frac{1}{ka}\right|_{f=f_o} = \left(\frac{\rho c^2}{3\gamma P_s}\right)^{1/2} \tag{5}$$

For bubbles sufficiently small to display natural frequencies in the kilohertz range, the quality factor is controlled by thermal damping. The pressure field, Eq. 5.26, is typically generated by pulsations excited transiently during bubble formation.

B. INTERACTION OF FLEXURAL WAVES IN A PLATE WITH AN AMBIENT FLUID

1. Velocity of Flexural Waves in an Elastic Plate In Vacuo

Consider the simple situation of an effectively infinite plate, conducting plane flexural waves. For statis loads, elementary bending theory yields

$$\frac{Eh^3}{12(1-\nu^2)}\frac{d^4w}{dx^4} = -p(x) \tag{6}$$

where E is the Young's modulus, ν the Poisson's ratio, and h the plate thickness. In the development below, the bending stiffness will be expressed in terms of the velocity of compressional waves in plates:

$$c_p = \left(\frac{E}{\rho_s(1-\nu^2)}\right)^{1/2} \tag{7}$$

where ρ_s is the density of the plate material. This velocity equals 5.42 x 10^5 cm/s for steel and aluminum plates. Under dynamic conditions, Eq. 6 must be modified to account for the inertia force per unit plate area. The equation of motion thus becomes:

$$\frac{h^3 \rho_s c_p^2}{12} \frac{\partial^4 w}{\partial x^2} - \rho_s h \omega^2 w = -p \tag{8}$$

Assuming plane or straight-crested flexural waves in the form of Eq. 5.15,

$$w = We^{\exp(ik_f x - i\omega t)}$$

the homogeneous equation governing free wave propagation is

$$h\rho_s \left(\frac{h^2 c_p^2 k_f^4}{12} - \omega^2 \right) w = 0$$

Solving for the flexural wavenumber,

$$k_f = 12^{1/4} \left(\frac{\omega}{h c_p} \right)^{1/2} \tag{9}$$

one finally obtains the flexural wave velocity

$$c_f = \frac{\omega}{k_f} = \frac{1}{12^{1/4}} (\omega h c_p)^{1/2} \tag{10}$$

In contrast to the sound velocity, the flexural wave velocity is dispersive.

2. Coincidence Frequency

We will see shortly that radiation loading depends critically on the magnitude of the ratio c_f/c. since c_f increases with $\omega^{1/2}$, it is less than the sound velocity at low frequencies and vice versa. The frequency at which the two velocities are equal is the coincidence frequency f_c. Equating c_f, Eq. 10, to c, and solving for frequency,

$$f_c = \frac{3^{1/2} c^2}{\pi h c_p} \tag{11}$$

For steel plates in air, f_c = 1,200/h and in water, f_c = 21,700/h, in Hz with h in cm. Actually, the product hf_c is sufficiently large for the latter expression to require corrections for rotary inertia and shear not accounted for in the simple Bernoulli-Euler theory in Eq. 8. Furthermore, as we shall shortly see, the flexural velocity is significantly reduced by radiation loading exerted by water. The crucial importance of the coincidence frequency was first recognized by Cremer (1942).

3. Pressure Radiated by Effectively Infinite Trains of Waves

Straight-crested flexural waves are matched to plane sound waves by the boundary condition in Eq. 5.16. The pressure is governed by the two-dimensional Helmholtz equation formulated in rectangular coordinates. Selecting a rectangular coordinate system whereby the z-axis is normal to the plate and the x-axis normal to the wavefront,

$$\frac{\partial p}{\partial z} = -\rho \ddot{W} e^{ik_s x} \quad ; \qquad z = 0$$

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} + k_s^2 p = 0$$

where k_s is the flexural wavenumber modified by radiation loading. The reader can verify that the pressure compatible with the above equations is of the form

$$p(x,z) = \frac{i\rho \ddot{W} \exp[ik_s x + i(k^2 - k_s^2)^{1/2} z]}{(k^2 - k_s^2)^{1/2}} \tag{12}$$

When $k > k_s$, i.e., for $f > f_c$, the denominator is real. The pressure is in phase with velocity, since

$$i\rho \ddot{W} = \rho c k \dot{W}$$

Radiation loading is resistive and tends to ρc when $k^2 >> k_s^2$. When $k < k_s$, the denominator is imaginary and the pressure in phase with acceleration. In this case, the reactive radiation loading is associated with the inertia force of the entrained mass of fluid. The mass per unit area is

$$m_s = \frac{\rho}{(k_s^2-k^2)^{1/2}} \quad , \quad k < k_s \quad , \quad f < f_c$$

$$\simeq \frac{\rho}{k_s} \quad , \quad k^2 << k_s^2 \quad , \quad f << f_c \tag{13}$$

Since the corresponding exponential in Eq. 12 is real, the pressure is in the form of non-propagating, exponentially decaying evanescent waves. When the latter of Eqs. 13 applies, the pressure field is associated with pseudo-sound such as observed in a strictly incompressible fluid:

$$p(x,z) \simeq \frac{\rho\ddot{w}}{k_s} \exp[k_s(ix-z)] \quad , \quad f << f_c \tag{14}$$

For a finite plate, or a point-excited effectively infinite plate (see Section C below), structural response is associated with a continuous wave-number spectrum, rather than with a single monochromatic spectral line. The supersonic portion of a continuous wavenumber spectrum will be shown to radiate sound even in the frequency range $f < f_c$.

4. Effect of Accession-to-Inertia on the Flexural Wave Velocity

The entrained mass of fluid Eq. 13, must be added to the mass $\rho_s h$ of the plate per unit area in Eq. 8. An asymptotic low-frequency expression for the structural wave-number is readily obtained when $k_s^2 >> k^2$, i.e., when $f << f_c$. The homogeneous equation of motion now takes the form:

$$\rho_s h\left[\frac{h^2 c_p^2 k^4}{12} - \left(1 + \frac{\rho}{\rho_s k_s h}\right)\omega^2\right] = 0 \quad , \quad k_s^2 >> k^2$$

If the plate is exposed to a fluid on both its surfaces, ρ must be multiplied by two. Since $k_s h$ becomes arbitrarily small with decreasing frequency, the unit term in the ω^2-coefficient can be neglected, and the wavenumber solved for. The real root associated with propagating waves is

$$k_s = \left(\frac{12\omega^2\rho}{h^3\rho_s c_p^2}\right)^{1/5} \quad , \quad k_s h << \frac{\rho}{\rho_s} \tag{15}$$

The ratio of the submerged to in vacuo flexural velocities and wavenumbers is

$$\frac{c_f}{c_s} = \frac{k_s}{k_f} = 0.883\left(\frac{c_p}{\omega h}\right)^{1/10}\left(\frac{\rho}{\rho_s}\right)^{1/5} \quad , \quad k_s h << \frac{\rho}{\rho_s} \tag{16}$$

The discrepancy between the two velocities is seen to increase with decreasing frequency.

5. Scaling Laws for Submerged Structures

The compatibility of the acoustic scaling laws formulated in Lecture 5 with structural and coupled structural-acoustic waves characterized by dispersion is not immediately apparent. Using the same acoustic fluid in the model test as in the full-scale situation, acoustical scaling laws require that if linear dimensions are divided by the scaling factor s, the model frequency be multiplied by s while velocities and pressures remain invariant. Retaining the same structural material in the model experiment, the flexural velocity in Eq. 10 is seen to remain unchanged, since ωh is constant. Using the same acoustic fluid in the model test, the structural wave number of the submerged plate, Eq. 15, varies as $(\omega^2/h^3)^{1/5}$. This factor, and therefore k_s, vary linearly with the scaling factor s, as does the in vacuo flexural wavenumber, Eq. 9. Consequently, the flexural wave vleocity remains invariant in vacuo as it does when radiation loaded. This is consistent with the fact that the accession-to-inertia per unit area, Eq. 13, varies as s^{-1}, as does the mass per unit area of the plate. In summary, as long as materials and acoustic media remain unchanged in the model experiment, acoustic scaling laws are compatible with the scaling of structure-borne noise and vibrations. In contrast to radiation damping, most types of structural damping are not properly modelled.

C. A FULL-FLEDGED INTERACTION PROBLEM: SOUND FIELD OF POINT-EXCITED PLATES

The sound radiated by a point-excited plate, radiation loading being taken into account is an important canonical problem solved by Feit (1966).[1] The mathematical model stipulates a plate whose extent, as measured in flexural wavelengths, insures that boundary reflections are negligible as a result of structural and radiation damping. This situation is properly modelled by an infinite plate. The equation of motion of the plate is in the form of Eq. 8, the differential operator being conveniently expressed in terms of cylindrical coordinates since the plate response and pressure field are symmetrical about the point of application of the concentrated force F. The z-axis is normal to the plate through this point, which is the origin of coordinates:

$$\frac{\rho_s c_p^2 h^3}{12}\left[\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}\right)^2 - k_f^4\right] w = -p(r,0) + \frac{F\delta(r)}{2\pi r} \tag{17}$$

Here $\delta(r)$ is the Dirac delta function. If the plate is radiation-loaded on both its surfaces the pressure term is multiplied by two. Applying the Hankel transform, Eq. 7.7, one obtains an algebraic equation governing the transformed plate response and surface pressure: We note that the transform of the differential operator is $(-\gamma^2)^2 = \gamma^4$. Furthermore, from Eq. 9, the bending stiffness can be written as $\omega^2\rho_s h/k_f^4$. The transform of Eq. 17 can now be stated:

$$\frac{\omega^2\rho_s h}{k_f^4}(\gamma^4-k_f^4)\tilde{w}(\gamma) = -\tilde{p}(\gamma,0) + \frac{F}{2\pi} \tag{18}$$

The pressure transform can be eliminated by means of the boundary condition, Eq. 7.10. Substituting the pressure transform, Eq. 7.10a, evaluated at z=0, and setting $\omega^2\tilde{w}(\gamma) = -\ddot{\tilde{w}}(\gamma)$, one obtains an equation which can be solved for the transform of the plate acceleration:

$$\ddot{\tilde{w}}(\gamma) = \frac{-F(k^2-\gamma^2)^{1/2}}{2\pi\{i\rho+(k^2-\gamma^2)^{1/2}[(\gamma/k_f)^4-1]\rho_s h\}} \tag{19}$$

There is no need to evaluate the inverse transform of the plate response if one is exclusively interested in the far-field pressure. The latter is formulated in terms of the stationary-phase value of the acceleration transform, Eq. 7.17. It is convenient to express the result in terms of the coincidence frequency, Eq. 11, by noting that

$$k^2/k_f^2 = f/f_c$$

Furthermore, the stationary-phase value of γ, Eq. 7.4, yields

$$(k^2-\gamma^2)^{1/2} = k\cos\theta$$

Substituting these results in Eq. 19, one obtains the stationary-phase value of the plate response transform. Combining this result with Eq. 7.17 the far-field pressure can be explicitly formulated:

$$p(R,\theta) = -\frac{iFk}{2\pi R}\,\frac{e^{ikR\cos\theta}}{1-ikh(\rho_s/\rho)\cos\theta[1-(f/f_c)^2\sin^4\theta]} \tag{20}$$

For $f/f_c < 0$, the peak occurs on the normal through the drivepoint:

$$p(R,0) = \frac{-iFk}{4\pi R}\,\frac{e^{ikR}}{1-ikh(\rho_s/\rho)} \tag{21}$$

Above coincidence, the peak occurs on the coincidence cone

$$p(R,\theta_c) = \frac{-iFke^{ikR}[1-(f_c/f)]^{1/2}}{2\pi R}\cdot\,, \qquad f > f_c \tag{22}$$

where

$$\theta_c = \sin^{-1}(f_c/f)^{1/2} = \sin^{-1}(k_f/k)$$

Plate parameters are embodied exclusively in the coincidence frequency.

REFERENCE

1 D. Feit, J. Acoust. Soc. Am. 40, 1489-1494 (1966).

THE APPLICATION OF STATISTICAL ENERGY ANALYSIS TO VIBRATION OF STRUCTURES EXCITED ACOUSTICALLY

B.L. Clarkson
Professor of Vibration Studies
Institute of Sound & Vibration Research
University of Southampton
Southampton SO9 5NH
England

SUMMARY

In the high frequency range of vibration which can be excited by acoustic pressures, there are usually very many modes of vibration. The individual mode by mode analysis becomes impracticable and an alternative approach based on the energy of vibration in broad frequency bands is more appropriate. By its very nature of averaging the effects over many modes the method is only suitable for randomly excited vibrations. This lecture gives an outline of the method and presents some experimental results for the parameters required.

LIST OF SYMBOLS

E	energy
F	force
f	frequency Hz
g	rate factor for energy flow
m	total mass of structure
n	modal density
P	power
t	time
v	velocity
Y	mobility (inverse of impedance)
Z	impedance
α,β	sub systems
η	loss factor
ω	frequency rads/sec

INTRODUCTION

In the previous lecture the response of a system has been expressed in terms of the responses in the individual modes of vibration of the structure. For strength and fatigue design considerations the low order modes usually give the highest stresses and so are the major contributors to the overall response. However, if we are interested in the vibration environment of equipment the acceleration of the support structure may be the most important design parameter. Stress is a function of displacement directly whereas acceleration is a function of displacement times the square of the frequency. Thus the high frequency region is often of importance when considering equipment vibration levels. In this high frequency range for many typical structures there are very many modes of vibration contributing to the overall vibration level. It now becomes impracticable to compute the responses in each of a large number of modes of vibration.

An alternative approach has been developed by Lyon (1) on the basis of some of the ideas used in room acoustics. The theoretical analysis of sound fields in rooms use probabilistic models (such as a diffuse sound field) and yield statistical measures of behaviour. The new method as applied to structural vibration has been called 'Statistical Energy Analysis' because energy is the independent dynamical variable which is used and it becomes statistical because only mean square values over fixed frequency ranges and averages over the surface of structural components are used. By considering energy flow it is possible to treat coupled mechanical systems, coupled fluid structural configurations, etc.

The aim of the analysis is to estimate the distribution of vibrational energy among the various elements of the coupled system. This is done by equating the power flowing into the system at the input point, to the power flowing from one subsystem to another plus the energy dissipated. Because of the statistical descriptions of the forces and system parameters, the equations involve average energies and power flows. These are expressed as averages over time and also over the modes of vibration having natural frequencies in a band of frequencies which is wide enough to include at least five modes. Because there will usually be some frequency dependence in the dynamic interactions the frequency bands should be chosen so as to highlight any such trends. The acousticians have established analysis procedures based on third octave bands but for many structural applications a 100 Hz band width might be more suitable.

The structures usually considered are finite linear elastic systems which can be described in terms of their uncoupled natural frequencies, mode shapes and dampings. Each normal mode is modelled as a simple oscillator and the interaction between two coupled subsystems can be represented by the coupling between two oscillator sets. The study of the response of coupled oscillators to random excitations lies at the heart of the method.

POWER INPUT AND DISSIPATION

The time averaged power flowing into a structure from a harmonic point force is given by:

$$P_i = \overline{F(t)\, v(t)}$$

$$= \frac{1}{2} |F_o|^2 \, \mathrm{Re}\,(Y)$$

Where F_o is the magnitude of the harmonic force and Y is the mobility of the structure at the point of application of the force. The power dissipated in a vibrating structure by internal losses etc. is given by:

$$P_d = \eta \omega_o E$$

where E is the time averaged energy of vibration and η is the loss factor.

POWER FLOW BETWEEN COUPLED OSCILLATORS

The power flow between two coupled oscillators which in general are each excited by broad band random forces can be obtained from the solution of the equations of motion of the two coupled systems (see Lyon (1) and Cremer, Heckle and Ungar (2)). The time averaged power flow from oscillator 1 to oscillator 2 is written in the form

$$\bar{P}_{12} = g(\bar{E}_1 - \bar{E}_2)$$

where $\bar{E}_1$ and $\bar{E}_2$ are the time averaged energies of the two systems. g is the rate factor.

The expression for g(1) (2) shows that energy flows on average from the oscillator with the higher energy to the one with the lower energy. The rate of flow is proportional to the energy difference. The rate factor g is a function of the system parameters and is independent of the excitation force levels. The rate factor is very sensitive to the differences in the individual 'blocked' natural frequencies of the two oscillators. Thus oscillators having close natural frequencies exchange energy more easily than those having well separated natural frequencies. For purely acoustic coupling the coupling element itself has no mass or stiffness and the 'blocked' natural frequencies are the same as the uncoupled natural frequencies of the systems.

COUPLING LOSS FACTOR

For ease in formulating the energy balance equations, it is convenient to express the power flow in terms of a coupling loss (analagous to dissipation loss). However as the power flow is dependent on the energy difference we must define the coupling loss factor η_{12} for the case of zero energy in system 2.

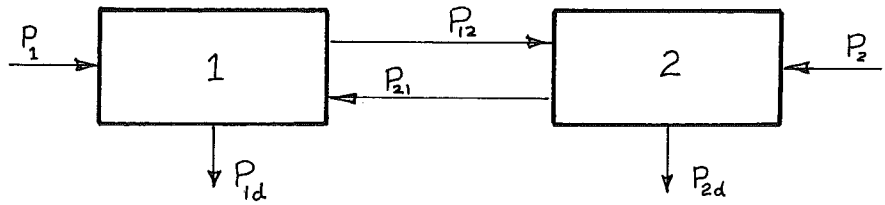
Thus:

$$\bar{P}_{12} = g\bar{E}_1 = \eta_{12} \omega_1 \bar{E}_1$$

$$\bar{P}_{21} = g\bar{E}_2 = \eta_{21} \omega_2 \bar{E}_2$$

$$\eta_{12} \omega_1 = \eta_{21} \omega_2$$

ENERGY BALANCE EQUATIONS



The energy balance equations for the two coupled systems illustrated above are:

$$\bar{P}_1 = \bar{P}_{1d} + \bar{P}_{12} - \bar{P}_{21} = \eta_1 \omega_1 \bar{E}_1 + \eta_{12} \omega_1 \bar{E}_1 - \eta_{21} \omega_2 \bar{E}_2$$

$$= \eta_1 \omega_1 \bar{E}_1 + \eta_{12} \omega_1 (\bar{E}_1 - \bar{E}_2)$$

$$\bar{P}_2 = \eta_2 \omega_2 \bar{E}_2 + \eta_{21} \omega_2 (\bar{E}_2 - \bar{E}_1)$$

If there is only one excitation force (say $\bar{P}_2 = 0$) we get:

$$\frac{\bar{E}_2}{\bar{E}_1} = \frac{\eta_{21}}{\eta_2 + \eta_{21}}$$

This shows that if the coupling loss factor is high compared with the internal dissipation the energies of the two systems will be approximately equal. This is known as strong coupling and equipartitioning of energy.

COUPLING OF MULTIMODAL SYSTEMS

If we now consider two multimodal systems (such as two beams or a beam and a plate) connected together we can extend the ideas discussed above to consider coupling between two modes of one system, and cross coupling between modes where one is chosen from each system. We now have to consider a large number of cross couplings but from our previous result we will only get significant energy flow when the frequencies of the two modes to be coupled are rather close together. We can now only consider groups of modes in

frequency bands (1/3 octave or 100 Hz) and work in terms of average coupling loss factors. The power flow equations must now be written in terms of averages over frequency bands and the coupling equation becomes:

$$\frac{\bar{E}_\beta}{\bar{E}_\alpha} = \frac{\eta_{\alpha\beta}}{\eta_{\beta\alpha} + \eta_\beta}$$

where η_β $\eta_{\alpha\beta}$ $\eta_{\beta\alpha}$ are frequency band averages and $\bar{E}_\alpha$ $\bar{E}_\beta$ are also frequency band average energies. For a uniformly distributed system such as a plate or a beam:

$$\bar{E}_\alpha = m_\alpha \langle \bar{v}_\alpha^2(t) \rangle$$

where m_α is the total mass of the plate or beam and $\langle \bar{v}_\alpha^2(t) \rangle$ is the spatial average of the time averaged square of the velocity.

Another important result relates the coupling loss factors to the modal densities of the two coupled systems:

$$n_\alpha \eta_{\alpha\beta} = n_\beta \eta_{\beta\alpha}$$

where n_α and n_β are the modal densities of system α and β respectively.

EVALUATION OF THE PARAMETERS REQUIRED

1. Modal Densities

Theoretical studies have produced results for the modal densities of many types of structure. Some of these are referred to in the bibliography (4). Some experimental results are also reported but the earlier experimental method of counting modes is rather inaccurate in all except the simplest cases. More recent work using an indirect method based on Cremer's results has produced experimental values for a range of structures. This result is:

$$n(f) = 4 M R_e (Y)$$

Thus frequency averaged values of the mobility can be used to derive the modal density. For example, the modal density of a cylindrical shell takes the form shown in figure 1.

2. Dissipation Loss Factor

Some theoretical results are available for the sound radiation component of dissipation loss in the high frequency range. However in general it must be admitted that we have to rely on experimentally derived results. These have been obtained by exciting the structure by a broad band of random force and then measuring the vibration decay rate when the force is cut off. This gives an estimate of the average loss factor for all the modes in the band. The method is satisfactory if all the modes have similar loss factors, but if one or two modes have much lower loss factors than the rest their slow decay will bias the average result. An example of the frequency band average loss factor for the cylindrical shell is shown in figure 2.

3. Coupling Loss Factor

Few results are available for structure to structure coupling. However there are several theoretical results for the coupling of an acoustic space to different structural configurations. These are given in the literature. An example of the coupling loss factor for a cylindrical shell to honeycomb platform function is shown in figure 3.

4. Input Power

The time average power input for a random force can be obtained by taking the frequency average of:

$$\overline{F^2(t)}\ R_e(Y)$$

The power input to a structure from an acoustic field can be obtained from the energy in the acoustic field and the coupling loss factor for the acoustic space and the particular structure.

REFERENCES

R.H. Lyon Statistical Energy Analysis of Dynamical Systems Theory and Applications. The MIT Press 1975.

L. Cremer, M. Heckl, E.E. Ungar Structure-Borne Sound. Springer-Verlag 1973.

F.J. Fahy Statistical Energy Analysis: A Critical Review. The Shock and Vibration Digest 1974.

F.D. Hart, K.C. Shah Compendium of Modal Densities of Structures. NASA CR-1773 1971.

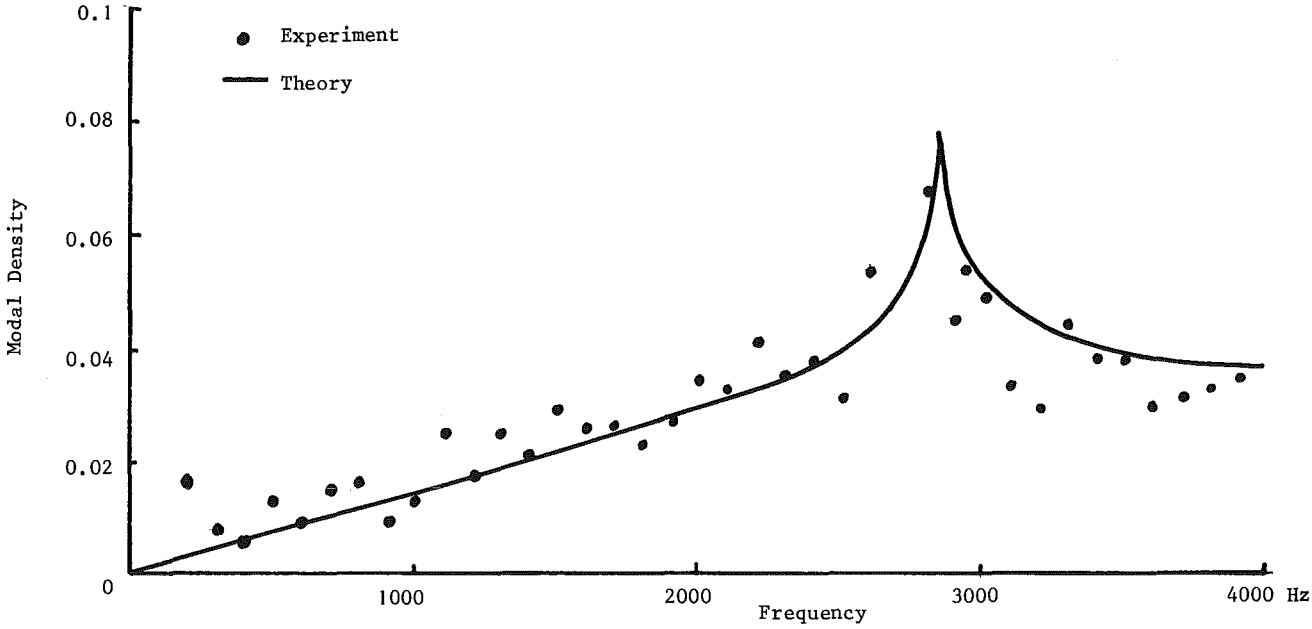


FIGURE 1 Modal Density of Cylindrical Shell (R = 0.3 m, ℓ = 0.75 m, t = 1.46 mm)

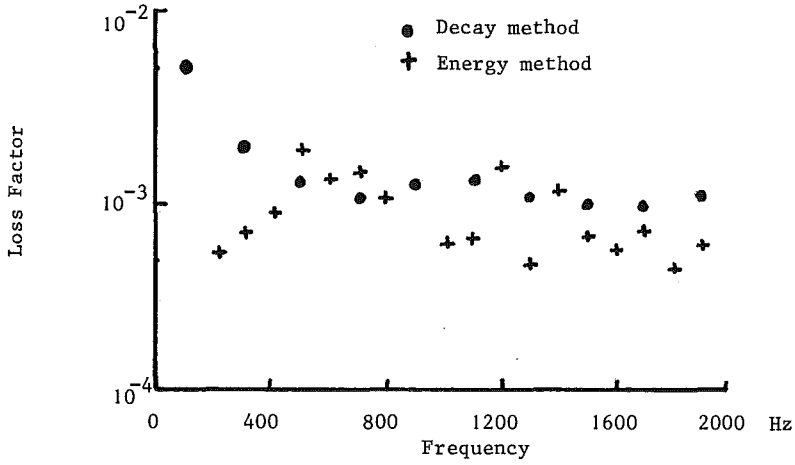


FIGURE 2 Loss Factor of Cylindrical Shell

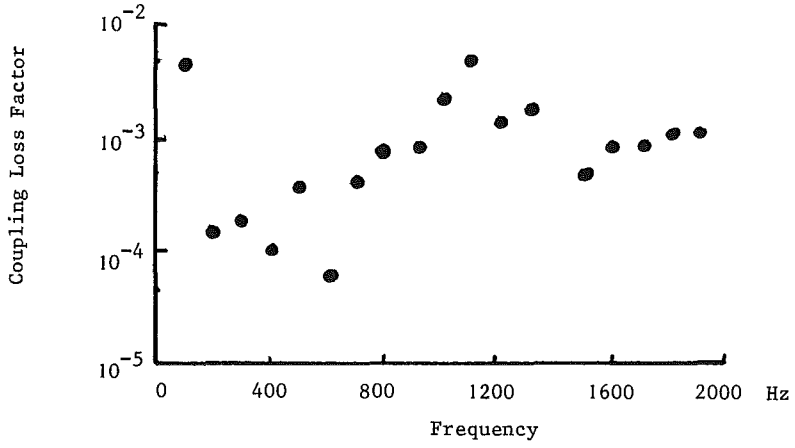


FIGURE 3 Coupling Loss Factor of Cylinder to Honeycomb Plate Junction

LINEAR SIGNAL PROCESSING I

J.K. Hammond
Institute of Sound and Vibration Research
University of Southampton
Southampton SO9 5NH
England

SUMMARY

This chapter is a review of methods of signal processing emphasising digital methods applicable to deterministic data. It is the first of two lectures devoted to basic procedures leaving more unusual and special problems (e.g., nonstationarity and nonlinearity) for separate treatment. Following a discussion of analogue to digital conversion, Fourier methods for continuous and sampled data are described, noting the problem of aliasing and the importance of circular convolution. Digital filtering techniques are referred to and methods of 'zoom' analysis are explained. The chapter concludes with a discussion of the problem of deconvolution.

1. INTRODUCTION

Data analysis involves the three phases of data acquisition, processing and interpretation (of the results of the processing) and of course all three are linked in any application. However, the last phase is naturally very much related to the particular phenomenon under investigation but the first two may be discussed independent of specific applications. A vast body of theory and methodology has been built up as a consequence of the problems raised by the need for data analysis often referred to as 'signal processing' or 'time series analysis' depending on the context. In four chapters on data analysis we shall attempt to give fairly comprehensive coverage of methods appropriate to deterministic data, random data and data having some particular structure (e.g., the output of a linear or nonlinear system).

This chapter is concerned with methods of analysis of time histories which have usually been obtained as measurements in the course of conducting an experiment, or when recording the behaviour of a physical system. The methods will be limited to 'basic procedures' (i.e., commonly used methods that are at the heart of signal processing and analysis). We will be concerned with deterministic data only, leaving the treatment of stationary random processes to the second chapter. More unusual methods of processing will be described in later chapters. In view of the fact that the audience may have differing backgrounds the treatment will stress fundamentals rather than points of detail. The references listed should serve to cover many specialist requirements.

Types of data to be analysed

Figure 1 broadly categorises the time histories that may arise [3]. A basic distinction is the designation of a signal as 'random' or 'deterministic', where by 'random' we mean one that is not predictable exactly. Very often processes are mixed and the demarcations of Fig. 1 are not easily applied and consequently the analysis procedure to be used may not be apparent. Even if we restrict data to the categories listed, the two classes of data called nonstationary random and almost periodic deterministic are often difficult to treat, and we shall discuss these in a later chapter.

Objectives of data analysis

Broadly speaking the objective of data analysis is to highlight/extract information contained in a signal that direct observation may not reveal. Sometimes only 'gross' characteristics of a signal may be required (e.g., mean value, mean square level, total energy) in which case some simple processing will provide the required quantities whilst ignoring other information contained in the signal which might only unnecessarily confuse matters. At other times it may be necessary to probe more deeply in an attempt to extract features of interest.

Methods of data analysis

The signals that are processed are time histories (often electrical outputs from transducers) or equivalent (e.g., rough ground height profiles where a space variable is the equivalent of time) and useful information extraction procedures may be performed whilst still in the 'time domain'. However, very often transform methods are employed since they offer simplifications both in theory and physical interpretation and 'frequency domain methods' are commonly used.

The signals recorded from physical phenomena are generally 'continuous time' in nature (sometimes called analogue data) but developments in digital computation methods have now ensured that most signal processing is achieved either on special purpose analysers or on larger (e.g., time shared, interactive) digital systems. This means the data is sampled and the analysis of discrete (in time) data requires that the analyst should have a good grasp of the fundamentals of digital signal processing, which is now a vast area of study (see, for example, [9, 10]). It is appropriate now to begin by briefly discussing the digitisation of continuous data.

2. ANALOGUE TO DIGITAL CONVERSION

An analogue-to-digital converter (ADC) is a device that operates on a continuous time history (input) and produces a sequence of numbers (output) that are sample values of the input. It is convenient to regard the process as consisting of two stages, namely, 'sampling' and 'quantisation' (see Fig. 2).

$x(n\Delta)$ is the exact value attained by time history $x(t)$ at time $t = n\Delta$ (Δ is the sample interval for uniform rate sampling). $\hat{x}(n\Delta)$ is the representation of $x(n\Delta)$ in a computer and differs from $x(n\Delta)$ since a finite number of bits are used to represent each number. Actual ADC's do not consist of two separate components (as in the 'conceptual' Figure 2) and a variety of types are available. One widely used type of ADC is that based on successive approximations (see Fig. 3). The digital word in the register is converted to an analogue signal and compared with $x(t)$. The error (which should be as small as possible) is used to update the digital word. The most significant bit in the register is adjusted first, working on down to the least significant. See refs. [1, 8] for additional details on ADCS.

The problem of quantization error is treated in [9,10,4,8]. Referring to Figure 2, the output $\tilde{x}(n\Delta)$ can be written as $\tilde{x}(n\Delta) = x(n\Delta) + e(n\Delta)$ and finite word length effects are described by treating $e(n\Delta)$ as 'noise'. The probability distributions describing the error depend on the particular way in which the quantisation is done. If the ADC introduces rounding (as opposed to truncation) error, then the error is ascribed a uniform distribution (with zero mean) over one quantization step and furthermore is assumed to be stationary and 'white'. For a b-bit word length (excluding sign), and if the range of the ADC corresponds to X volts, say, then the variance of the error $\sigma_e^2 = (X/2^b)^2/12$. Furthermore if x is assumed to be a random signal (see [9]) and σ_x^2 denotes the variance of $x(n\Delta)$, then a measure of signal-to-noise ratio (on a logarithmic scale) is $S/N = 10 \log_{10}(\sigma_x^2/\sigma_e^2)$. If $\sigma_x = X/4$ (to minimise clipping) then $S/N = 6b - 1.24$ dB giving a value of about 70 dB for a 12 bit ADC. This is reduced further by practical considerations (see [4]).

The rate at which $x(t)$ is sampled should be high enough to avoid aliasing (see section 4). This means that $x(t)$ should usually be low pass filtered using analogue filters preceding the ADC. The cut-off frequency and cut-off rate of these 'anti-alias' filters should be chosen with particular applications in mind but very roughly speaking if the 3 dB point is a quarter of the sampling frequency and the cut-off rate better than 48 dB/octave, then this at least ensures a 40-50 dB reduction at the folding frequency (half the sampling frequency), and so probably an acceptable level of aliasing (though we must emphasise that this may not be adequate for some applications).

3. DETERMINISTIC DATA - FOURIER METHODS

In Figure 1 we defined three types of deterministic signal, namely, periodic, transient and almost periodic. Periodic data is analysed using the Fourier series, transient data uses the Fourier integral and some classes of almost periodic signals (and also damped signals) may be studied using the Prony series [15]. The basic difference between the Prony series and the Fourier series is that in the former case the frequencies are not assumed known and so both amplitude, phase and frequency are computed, while only amplitudes and phase are calculated at particular frequencies using the Fourier method. The Prony method is computationally more 'difficult' but can be very attractive under certain circumstances (e.g., where data lengths are short). We shall only discuss Fourier methods in this section.

3.1 The Fourier Series and Fourier Integral

If a signal repeats itself exactly every T_p seconds then $x(t)$ may be represented as

$$x(t) = \sum_{n=-\infty}^{\infty} C_n e^{j2\pi nt/T_p} \tag{1}$$

whilst if $x(t)$ is a transient then the discrete set of sines and cosines becomes a continuum and we write

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df. \tag{2}$$

The specification of C_n or $X(f)$ is equivalent to the original time history. For example, $X(f)$ is given by

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt. \tag{3}$$

An important consideration when calculating the transform $X(f)$ is that of data truncation (or windowing). The infinite integral (3) may be replaced by a finite one because, for instance, $x(t)$ may only be known (measured) for a limited duration, say for T seconds, and we might define $x_T(t) = x(t)$, $|t| \leq T/2$; $= 0$, elsewhere. x_T is a truncated version of $x(t)$ and from Figure 4 we can think of this cutting off the tails of the data as though we are looking at the data through a 'data window' $w(t)$, where

$$w(t) = 1, \quad |t| \leq T/2$$
$$= 0, \quad \text{elsewhere},$$

so that

$$x_T(t) = x(t)w(t).$$

Fourier transforming $x_T(t)$ gives $X_T(f)$ which is related to the Fourier transforms of x and w by

$$X_T(f) = \int_{-\infty}^{\infty} X(g)W(f - g)\,dg \tag{4}$$

where in turn $W(f)$ is sometimes called a 'spectral window'.

The convolution integral (4) indicates that $X_T(f)$ is a distorted version of $X(f)$ (unless $W(f)$ is a delta function). This effect is simply demonstrated with a truncated cosine wave. If $x(t) = \cos 2\pi pt$, then

$$X(f) = \tfrac{1}{2}[\delta(f + p) + \delta(f - p)] \qquad \text{(Fig. 5(a))}$$

and

$$X_T(f) = \tfrac{1}{2}[W(f + p) + W(f - p)] \qquad \text{(Fig. 5(b) for a rectangular data window)}$$

The appearance of energy at frequencies other than p is sometimes called leakage. If two frequencies are present, say p_1, p_2, then in order to be able to get two distinguishable peaks it is necessary for the data length T to (approximately) satisfy $T \geq 2/(p_2 - p_1)$ for a rectangular window. The rectangular window is a poor window since the side lobes are large and do not decay rapidly. Numerous alternatives have been suggested to reduce leakage in Fourier transform calculations. By tapering the windows to zero, the side lobes are significantly reduced but at the expense of widening the main lobe. The frequency

domain behaviour of windows is often given in decibels (dB) where the characteristics have been normalised to unity gain (0 dB) at zero frequency and some terms are defined in Figure 6 which relates to the rectangular window.

The paper by Harris [11] contains a comprehensive treatment of windows and their effects.

3.2 The Fourier Transform of a Sequence

Since the calculations of Fourier coefficients and transforms are carried out on a digital computer it is necessary to re-examine the preceding work for data in discrete form. We might treat a discrete version of (3), for example, as an approximation to the 'correct' result. However completely rigorous, exact theories can be developed for discrete data which are, of course, closely related to the continuous time counterparts, but there are fundamental differences and it is often more convenient to consider the problems of discrete data as self-contained and relate the results to those of the continuous domain only as an when they are required, accounting at that stage for any errors or approximations that might have been incurred in using discrete methods in place of the continuous operations. Having said this we will, nevertheless, attempt to relate the Fourier analysis of continuous and discrete data.

Impulse train modulation

One way of introducing the Fourier transform of a sequence involves the use of the mathematical notion of 'ideal sampling' of a continuous wave. If an analogue signal $x(t)$ is to be sampled every Δ seconds it is convenient to model the sampled signal as the product of the continuous signal with a 'train' of delta functions $i(t)$ where

$$i(t) = \sum_{n=-\infty}^{\infty} \delta(t - n), \qquad \text{i.e.,} \quad x_s(t) = x(t)i(t).$$

The Fourier transform of $x_s(t)$ is $X_s(f)$ and using properties of the delta function, it is written

$$X_s(f) = \sum_{n=-\infty}^{\infty} x(n\Delta)e^{-j2\pi fn\Delta}. \tag{5}$$

This has the inverse

$$x(n\Delta) = \Delta \int_{-1/2\Delta}^{1/2\Delta} X_s(f)e^{j2\pi fn\Delta}df. \tag{6}$$

Equations (5) and (6) relate the *sequence of numbers* $x(n\Delta)$ to the quantity $X_s(f)$ which is termed the Fourier transform of the sequence. A natural question to ask at this stage is 'How is $X_s(f)$ (or as it is often written $X(e^{j2\pi f\Delta})$) related to $X(f)$?' The answer to this can be obtained formally as follows. Since $i(t)$ is periodic we might represent it with a Fourier series, i.e.,

$$i(t) = \frac{1}{\Delta} \sum_{n=-\infty}^{\infty} e^{(2\pi jnt)/\Delta}$$

Now substituting this into $x_s(t) = x(t)i(t)$ and Fourier transforming, gives an alternative right hand side to (5), namely,

$$X(e^{j2\pi f\Delta}) = \frac{1}{\Delta} \sum_{n=-\infty}^{\infty} X(f - \tfrac{n}{\Delta}). \tag{7}$$

This important equation relates the Fourier transform of a continuous signal and the Fourier transform of the sequence formed by sampling the signal at equispaced intervals. It will be discussed further in the section on 'aliasing'.

The z-transform

An alternative route to equation (5) is to use the notion of the z-transform of a sequence (widely used in the solution of difference equations). The z-transform of a sequence of numbers $x(n)$, say (where the notion of time is not made explicit) is

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} \tag{8}$$

z is interpreted as a complex number and $X(z)$ a function of a complex variable. If z is allowed to take values on the unit circle in the z plane, i.e., $z = e^{j\omega} = e^{j2\pi f}$, then (8) is

$$X(e^{j2\pi f}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j2\pi fn}$$

which is seen to correspond identically with (5) if the sample interval Δ is 1. The general inverse of (8) is

$$x(n) = \frac{1}{2\pi j} \oint_c X(z)z^{n-1}dz \tag{9}$$

where the contour c includes the singularities of $X(z)$, and which, if $x(n)$ is a bounded sequence, can be taken as the unit circle so that

$$x(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} X(e^{j2\pi f})e^{j2\pi fn}df. \tag{10}$$

4. ALIASING

Equation (7) describes how the frequency components of the sampled waveform relate to that of the continuous waveform and Figure 7 explains this pictorially. In Figure 7(a) it is assumed that $X(f) = 0$ for $|f| > f_o$ (say) and Figure 7(b) is a plot of $X(e^{j2\pi f\Delta})$ assuming that the sampling rate $f_s = 1/\Delta$ is such that $f_o < 1/2\Delta$. Some commonly used terms are defined on the diagram.

Suppose now $f_o > f_s/2$, then there is an overlapping of the shifted versions of $X(f)$ resulting in a distortion of the frequency description for $|f| < 1/2\Delta$ as in Figure 7(c). This 'distortion' is due to the fact that high frequencies in the data are indistinguishable from lower frequencies owing to the sampling rate not being fast enough. More specifically consider $x(t) = \cos 2\pi pt$ sampled at $f_s = 1/\Delta$, i.e., at $t = n\Delta$ (with, say, $p < 1/2\Delta$). This may be written $x(n\Delta) = \frac{1}{2}(e^{j2\pi n\Delta p} + e^{-j2\pi n\Delta p})$. Now letting frequency p above be replaced by $\pm p + k/\Delta$ $(k = 1, 2, \ldots,)$ results in an identical sequence. So if a frequency component were detected at p Hz, any one of these higher frequencies might have been responsible for this rather than a 'true' component at p Hz. This phenomenon is called 'aliasing'.

To avoid aliasing the sample rate must be chosen to be greater than twice the highest frequency contained in the signal. (See the discussion in Section 2 regarding 'anti-alias' filters.)

5. THE DISCRETE FOURIER TRANSFORM

The equation (5) is still not a 'practical proposition' and a finite sum would be more appropriate which, for an N point sequence, could be written

$$\sum_{n=0}^{N-1} x(n\Delta)e^{-j2\pi fn\Delta}.$$

Obviously, if this involved truncating the data some distortion would be involved, explained by the discussion on windowing earlier. Assuming for the moment that we can ignore this, we can think of simplifying our task by only evaluating f at specific points, say at $k/N\Delta$ Hz, $k = 0, \ldots, N-1$. Then

$$X(e^{\frac{j2\pi k}{N}}) = \sum_{n=0}^{N-1} x(n\Delta)e^{-j(\frac{2\pi}{N})nk}.$$

Or, adopting the more usual notation (and omitting the Δ for convenience),

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j(\frac{2\pi}{N})nk}. \tag{11}$$

Multiplying both sides by $e^{j(2\pi/N)mk}$ and summing over k gives

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k)e^{j(\frac{2\pi}{N})nk}. \tag{12}$$

The pair (9) and (10) constitute the Discrete Fourier Transform (DFT) and are the form that are suitable for machine computation. It is important to realise that even though the original sequence $x(n\Delta)$ may have been zero for n outside the range $0 \to N-1$, the act of 'sampling in frequency' has imposed a *periodic structure* to the sequence, i.e., it follows from (12) that $x(n + N) = x(n)$. It is instructive to follow (diagrammatically) the interpretation of a DFT of an aperiodic function that is truncated. (See Figure 8 (ref. [7]).)

The evaluation of the DFT can be accomplished efficiently by a set of algorithms known as the fast Fourier transform (FFT). The FFT algorithms essentially exploit the periodicity and symmetry properties of $e^{j(2\pi/N)nk}$, and so reduce the number of multiply and add operations needed to calculate $X(k)$ from about N^2 to approximately $N \log_2 N$ which for $N = 1024$ is a reduction by a factor of about 100. References [9] and [10] treat the subject comprehensively.

5.1 Circular convolution

A key operation in linear system theory is that of convolution. If $x(t)$ is an input and $h(\tau)$ the response of a linear time invariant system to a unit impulse then the output $y(t)$ may be expressed

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau. \tag{13}$$

If the system is causal, the lower limit is replaced by 0. The frequency domain equivalent of (13) obtained by Fourier transforming the equation is

$$Y(f) = H(f)X(f) \tag{14}$$

which is a considerable simplification (and from which we directly obtain a relationship between input and output energy spectra).

When discrete data are used the DFT is often used as a quick method (using the Fast Fourier Transform) to perform convolutions and in this context the notion of circular convolution and the 'wraparound' effect arises. We start by considering two sequences $x(n)$ and $h(n)$ defined for $0 \leq n \leq N-1$. Since we shall be considering DFT's of these sequences it is convenient to consider $x(n)$ and $h(n)$ as one period of periodic sequences $x_p(n)$, $h_p(n)$. For such periodic sequences Figure 9 explains the meaning of $x_p(n - n_o)$.

$x_p(n)$ may be considered as formed from observing a rotating drum with $x(n)$ on its circumference (hence the term 'circular').

If another periodic sequence is formed from a convolution as

$$y_p(n) = \sum_{\ell=0}^{N-1} x_p(\ell)h_p(n - \ell) \tag{15}$$

then

$$Y_p(k) = H_p(k)X_p(k). \tag{16}$$

The inverse discrete Fourier transform (IDFT) of the product of the two DFT's is a periodic sequence as in (15). This must be borne in mind when using the DFT to perform the convolution of two *non periodic* sequences $x(n)$, $h(n)$. If $x(n)$ is length N_1, and $h(n)$ length N_2, then

$$y(n) = \sum_{k=0}^{n} h(n - k)x(k)$$

is of length $N_1 + N_2 - 1$. If one selected the larger of the two sequences (say N_1) and performed N_1 point DFT's of both x and h (suitably augmenting the other sequence with zeros) and took the DFT of the product, the resulting sequence, $y_1(n)$ say, would be incorrect owing to the fact that it would be the result of convolving two *periodic* N_1 point sequences. To ensure that the correct result is obtained (within one period) it is necessary to augment both x and h with zeros to make them $N_1 + N_2 - 1$ point sequences. Then the IDFT of the product of DFT's of these sequences produces a periodic sequence, one period of which is the required result. This idea is basic to various methods of 'fast convolution' that make use of the fast Fourier transform.

6. DIGITAL FILTERING OF DATA

The computer offers enormous flexibility in the processing of data, and some of the vast array of procedures that may be carried out on a sequence of numbers is referred to as 'digital filtering'. Several extremely good books exist [9, 10, 8] that have become standard texts and all we shall do in this short section is introduce some terminology and direct the reader toward some of the methods that may prove relevant. We shall restrict our discussion to linear operations on input data with time (shift) invariant filters.

A linear time invariant system can be characterised by the impulse response sequence $h(n\Delta)$ (the response to a sequence which is zero everywhere except at $n = 0$, where it is unity). Then the response $y(n\Delta)$ to input $x(n\Delta)$ can be written (for a causal system)

$$y(n\Delta) = \sum_{m=-\infty}^{n} h((n - m)\Delta)x(m\Delta)$$

from which the frequency response function for the system is

$$H(e^{j2\pi f\Delta}) = \sum_{r=0}^{\infty} h(r\Delta)e^{-j2\pi fr\Delta}$$

which differs from the frequency response function of continuous systems in that it is a *periodic* function of frequency.

Rather than characterising the system (digital filter) by its impulse response sequence a difference equation form can be used to relate input and output sequences. More specifically, if $x(n)$ denotes an input sequence (time increment Δ is dropped for convenience), and $y(n)$ the output sequence, then a difference equation relating the sequences can be written

$$y(n) = -b_1y(n - 1) - b_2y(n - 2) \ldots -b_Ny(n-N) + a_ox(n) + a_1x(n - 1) + \ldots + a_Nx(n-M) \tag{17}$$

This is a general digital filter which can be programmed to produce an output sequence given an input sequence and some starting conditions. The z-transform of section 3 may also be used to solve the equation, and this leads to a definition of a transfer function, $H(z)$, where

$$H(z) = \frac{a_o + a_1z^{-1} + a_2z^{-2} + \ldots + a_Mz^{-M}}{1 + b_1z^{-1} + \ldots\ldots + b_Nz^{-N}} . \tag{18}$$

By suitable selection of the coefficients b_i, a_i (and the orders N, M) the transfer characteristics can be adjusted to approximate some desired form. Finite word length effects mean that ideal coefficients cannot be represented exactly and arithmetic round-off in the multiply operations constitute another source of error. These effects are not discussed further here (see chapter 9, ref. [9], for example, for more details).

Returning to equation (17), if at least one coefficient b_i is not zero the filter is recursive whilst if all the b_i are zero then the filter is non-recursive. If the filter has a finite memory (i.e., its impulse response sequence is zero after a certain 'time') then it is called a Finite Impulse Response or FIR filter, in contrast to Infinite Impulse Response (IIR) filters which have infinite memory. We note that the terms recursive and non-recursive describe how a filter is realised and not whether an impulse response is finite or infinite, though FIR filters are usually non-recursive and IIR filters recursive in implementation.

Numerous design procedures are available for both types of filter. A popular procedure for IIR filters is to discretize some well understood analogue filter (bearing in mind that any digital counterpart is bound to have a frequency response that is periodic in frequency). The 'impulse invariant' method of design, as its name suggests, creates a filter whose impulse response sequence matches the impulse response function of an analogue filter at equispaced time instants. This procedure is simple but limited in effectiveness since high pass and bandstop filters cannot be designed this way. The 'bilinear transform' method is a procedure that overcomes the 'aliasing' problems of the impulse invariance procedure at the expense of introducing some frequency distortion when transforming from the analogue to the digital domain. However, this can be compensated for and this is an effective and widely used procedure. Finally, optimal IIR filters can be designed to meet certain criteria and listings of Fortran programs are available [16], which produce optimal filters.

FIR filters have certain advantages over IIR filters, for example, that they are always stable and can have linear phase characteristics, but a major disadvantage is that they must be long enough to achieve adequate cut-off. Easily implemented Fortran FIR optimal filter design programs (e.g. [13]) are available that work very effectively, and such filters could then be implemented using a (fast) convolution algorithm. Recent technological developments have exploited the use of FIR filters (or tapped delay lines as they are often called) and filtering, correlation, etc., can be implemented cheaply and simply using CCD's (Charge Coupled Devices). A basic text covering much of this subject is [23].

We note that linear phase FIR filters can easily be designed (thereby ensuring no phase distortion), but, if data files can be stored then zero phase filtering can be achieved even with IIR filters. This is done by processing the data 'forwards' and 'backwards' with the same filter. Figure 10 shows a possible scheme (see [10]). The only thing to watch for are the filter 'starting transients' at both ends of the data, but this apart it is a simple and effective procedure.

7. 'ZOOM' ANALYSIS

Very often there is a need to study the 'fine detail' in a spectrum, i.e., to select a particular narrow band of the spectrum whose spectral lines are spaced at $1/N\Delta$ Hz apart ($f_s = 1/\Delta$ being the sample rate and N the number of data points in the FFT), and 'zoom' onto this band to obtain a finer spacing of points on the frequency axis. This is often referred to as 'improving the resolution', though we make the point that it is probably safer to refer to it as simply achieving a finer spectral line spacing, since true resolution may already have been limited by data truncation (prior to any transformation) so that finer spacing in such a case will merely give a closer look at a windowed spectrum. Bearing this in mind we note that there are several ways of achieving this finer spectral spacing and we shall discuss two of them.

7.1 The Chirp z-transformation Algorithm

The Chirp z-transformation (CZT) algorithm [9] is a procedure for the efficient computation of the z-transform $X(z)$ of a sequence $x(n)$, where

$$X(z) = \sum_{n=0}^{N-1} x(n)z^{-n}.$$

The ordinary DFT evaluates $X(z)$ at $z = e^{j(2\pi/N)k}$ and 'zoom' analysis requires the spacing to be finer than this over a restricted range. The CZT is in fact more general in that it is a procedure for the computation of the z-transform on a spiral contour in the z-plane at points that are equi-spaced in an angular sense (see Fig. 11).

If the points in the z-plane are M points, $z_k = AW^{-k}$ $(k = 0, \ldots, M-1)$, where $W = W_o e^{-j\phi_o}$ and $A = A_o e^{j\theta_o}$, (W_o, A_o real), then the spiral shown in the figure is defined (for $W_o < 1$) and the z-transform is

$$X(z_k) = \sum_{n=0}^{N-1} x(n)A^{-n}W^{nk} \qquad k = 0, \ldots, M-1.$$

Now employing the identity $nk = \frac{1}{2}[n^2 + k^2 - (k-n)^2]$ converts the required evaluation into a convolution (of complex valued sequences) which can be evaluated using 'fast convolution'. For 'zoom' analyses on the unit circle the parameters A_o and W_o would be set to unity. Whether this procedure is more efficient than a direct evaluation of the transform depends on the sizes of M and N (neither of which need be a power of two). There are other methods that may be employed for restricted evaluations on the unit disc (see, for example, the Goertzel algorithm in ref. [9]).

7.2 Zoom Analysis by Complex Demodulation (Heterodyning)

The key point to note in the frequency analysis of a time history using straightforward DFT (FFT) methods is that line spacing is at f_s/N where f_s is sample frequency and N is FFT size. If N is fixed (e.g., 1024) then f_s must be reduced, though obviously this cannot be done arbitrarily since aliasing might occur. The reduction in f_s can be achieved by first employing complex demodulation (often referred to as heterodyning, ref. [5]). We shall explain the procedure starting with a continuous time signal.

Suppose $x(t)$ has Fourier transform $X(f)$ (see Fig. 12a) with upper frequency f_{max}, and that the signal is sampled at f_s ($\geq 2f_{max}$) then suppose that if the resulting 'usual' spacing f_s/N is inadequate and we wish to obtain finer resolution in the shaded band. If we form $x(t)e^{-j2\pi f_c t} = x_1(t)$ we then create a complex valued signal $x_1(t)$ by a process called complex demodulation (or heterodyning) whose Fourier transform $X_1(f)$ has the form in Fig. 12b, i.e., a frequency shifted version of $X(f)$. If both components of $x_1(t)$ are now low pass filtered through a filter with sharp cut-off and bandwidth $B/2$, we obtain a narrow band process that can be sampled at f_s' ($< f_s$) and satisfying a much less stringent condition than $f_s \geq 2f_{max}$ so that the resolution using FFT

methods if f_s'/N.

If, however, the data has already been sampled at f_s and as a result of a preliminary spectral analysis it is decided that 'zooming' onto a particular frequency range is wanted, then the same technique may be employed by multiplying by

$$e^{-j2\pi f_c \frac{n}{f_s}} \qquad (n = 0, 1, \ldots).$$

However, owing to the fact that the data has already been sampled, the digital low pass filtering operation that is now necessary must be performed with care. Assuming that this is done without undue difficulty (difficulties may occur if we wish to zoom onto a region very near the folding frequency $f_s/2$), then the sample rate f_s is now unnecessarily high and so can be reduced. Various digital filtering methods exist for sample rate reduction but a simple and effective procedure is simply to retain only every p^{th} sample if a new sample rate of f_s/p is wanted (i.e., a zoom factor of p). The frequency analysis of the demodulated, decimated data then proceeds normally. It is worth noting that if a fixed record length of data has been acquired say for random data, then the extra length of data taken for each transformation means less smoothing is achieved if the segment averaging technique is used (see the next chapter).

The instantaneous envelope and phase

In view of the use of the complex demodulation method of 'zoom' analysis, it is appropriate to note that this technique can offer additional information about band limited signals if viewed slightly differently. It stems from the fact that any real band limited signal can be expressed in the form $x(t) = A(t) \cos [g(t)]$ where $A(t)$ is the envelope of $x(t)$ and $g(t)$ the phase. A unique definition of $A(t)$ and $g(t)$ can be obtained by using concepts from communication theory [25]. Suppose a band limited signal $x(t)$ has the spectrum shown in Fig. 13(a) then complex demodulation and filtering (b), and subsequent multiplication by $e^{j2\pi f_c t}$ results in a complex valued process $x^+(t)$ whose spectrum $X^+(f)$ is zero for $f < 0$ and identical to that of $2X(f)$ for $f > 0$. $x^+(t)$ is called the *analytic signal* associated with $x(t)$ and can be written

$$x^+(t) = A(t)e^{jg(t)} = x(t) + jx_h(t) \qquad (19)$$

so that the envelope $A(t)$ is $|x^+(t)|$ and the phase is $\arg x^+(t)$. The real valued quantity $x_h(t)$ is is the Hilbert transform of $x(t)$ and can be computed in a quite different way as the convolution of $x(t)$ with $1/\pi t$.

8. DECONVOLUTION

Deconvolution is a term which embraces a wide range of signal processing theory. The aim is to unravel the effects of a convolution of two signals (or a system and signal) so as to restore one of the signals. For example, if a signal x is operated upon by a system with impulse response h to produce an output y, we seek to design and apply some operator f which will restore x.

It is assumed that the operation on x is linear and that h and y are known. The operator f is the 'inverse filter' corresponding to h and may be denoted h^{-1}. However, it may not always be possible to realise an exact inverse, in which case it is necessary to make do with an approximation to h^{-1} and a corresponding approximation to the signal x. We therefore write the output of f as $\hat{x}$ (the 'hat' notation denotes an estimate of x) and define an error $e = x - \hat{x}$. We now try and find the filter f that minimises some function of the error. This is represented in figure 14. The problem posed relates to the deconvolution problem in its barest form. More usually the measurement y is noise contaminated and/or the system dynamics h are not known exactly (indeed the system may be nonlinear). It is these embellishments to the basic theme which account for the large number of techniques that have been developed with the objective of deconvolution.

It is appropriate now to point out a generalisation which is also sometimes referred to as deconvolution. This is the design of some filter which shapes a given (measured) signal y into some other desired form denoted by d. At first sight this problem appears unconnected with deconvolution. However, the ideal solution is $f = y^{-1} * d$, thus requiring the 'inversion' (deconvolution) of a signal. A special case of this is the predictive filter where the aim is to design a filter whose output is a future (time shifted) version of the input. In this section we shall mention some of the many approaches to the problem using both continuous time and discrete time formulations.

8.1 'Direct' Inversion

We shall look at the simplest case here, namely when measurement y is noise-free and h is known exactly. The treatment will be in discrete time. (* will denote convolution.)

Inverse digital filtering

Suppose that h is a discrete time system with known transfer function $H(z)$ which may be expressed as

$$H(z) = \frac{b_o + b_1 z^{-1} + \ldots b_M z^{-M}}{a_o + a_1 z^{-1} + \ldots a_N z^{-N}} = \frac{b_o \prod_{i=0}^{M} (1 + \alpha_i z^{-1})}{a_o \prod_{j=1}^{N} (1 + \beta_j z^{-1})}$$

with zeros at α_i and poles at β_j (the poles β_j satisfy $|\beta_j| < 1$ for stability). Now we seek f such that $h(n) * f(n) = \delta(n)$ ($\delta(n)$ denoting a unit impulse) or equivalently $H(z)F(z) = 1$, and we write, formally, $F(z) = 1/H(z)$. Unfortunately of course such a solution is not generally possible since the zeros of $H(z)$ need not lie within the unit disc in the z-plane (i.e., $H(z)$ may not be

minimum phase). Thus $1/H(z)$ will have poles outside the unit disc and will therefore be unstable if realised as a difference equation in 'forward time'.

Three possibilities now present themselves:-

(a) We form the inverse of the minimum phase version of $H(z)$. More specifically, any general transfer function $H(z)$ may be written (see reference [9]) $H(z) = H_{eq}(z)\, H_{ap}(z)$. $H_{eq}(z)$ denotes a transfer function having identical modulus to $H(z)$ but with zero locations adjusted to be within $|z| = 1$, and $H_{ap}(z)$ is an 'all pass' transfer function having unity modulus but with zeros outside the unit disc.

The all pass component has zeros and poles at conjugate reciprocal locations. The portion $H_{eq}(z)$ will have a stable, realisable inverse in forward time and so the operations indicated in fig. 15 will result in a signal $\hat{x}$ that is distorted by the all pass component. The *modulus* of the Fourier transforms of x and $\hat{x}$ are indistinguishable but the phase structures will differ.

(b) If we do not restrict ourselves to processing only in forward time then a factorisation of $H(z)$ into the form $H(z) = H_{min}(z)H_{max}(z)$ where $H_{min}(z)$ denotes the minimum phase part of $H(z)$ and $H_{max}(z)$ the maximum phase part of $H(z)$ permits complete deconvolution. This is demonstrated for a single zero lying outside the unit disc. Suppose we have $H(z) = 1 + kz^{-1}$, $|k| > 1$.

Clearly, $1/H(z)$ is unstable in forward time. However, it can be argued that the inverse is stable in *reverse* time and may be written as

$$\frac{\frac{1}{k}z}{1 + \frac{1}{k}z}\,.$$

(c) The third possibility is simply an approximate version of (b). If it is imperative that processing is done in forward time then the scheme (b) can be modified by calculating the impulse response function of the 'reverse time inverse' which is 'acausal' as far as the original data is concerned, truncating it at some suitable point, delaying it by its duration (to make it causal) and implementing it as a finite impulse response filter. The delay imposes a linear phase structure to the output but more importantly the truncation introduces distortion to amplitude and phase. This problem is addressed by Robinson (reference [17]), where a least squares approach is used to aid the design of the inverse, (see below).

As a final discussion on 'direct' methods we consider the 'Fourier' procedure of undoing $h(n)$ by dividing the Fourier transform of y by the Fourier transform of h and inverse transforming the result. This method follows from the discussion in section 5. Recall that if $x(n)$ is an N_1 point sequence, $h(n)$ an N_2 point sequence, then

$$y(n) = h(n) * x(n) = \sum_{m=0}^{n} h(m)x(n - m)$$

is an $L = N_1 + N_2 - 1$ point sequence. To obtain this convolution using the discrete Fourier transform (FFT) and to avoid distortions in the result due to 'circular convolution' zeros must be appended to $x(n)$ and $h(n)$ to make both say P point sequences ($P \geq L$). Then the product of the P point DFT of $h(n)$ and $x(n)$ gives the DFT of a periodic sequence, one period of which is the required sequence $y(n)$. So we can write $X(k) = Y(k)/H(k)$.
N.B. It is necessary, however, to resist the temptation to say that the sequence $h_1(n)$ whose DFT is $1/H(k)$ is the impulse response function of the inverse filter of $h(n)$. Oppenheim and Schafer give an instructive example on page 132 of reference [9] to clarify this point.

We note in conclusion that these discussions, relating as they do to ideal inversion, take no account of additive noise. It is worth noting that a compromise between inverse filter bandwidth and output signal to noise ratio may be considered by relaxing the requirement that $h * f = \delta$ and instead of the delta function use a more 'spread' function.

8.2 Optimal Filters

When it is not possible to form an ideal inverse we are obliged to construct an approximation. One particularly useful and widely used procedure is to design the filter that minimises some squared error. Referring to figure 16, we seek f that minimises some measure of e^2. There are several versions of this problem and its solution and we shall note a few of them briefly. For some sort of completeness both continuous and discrete time formulations are considered. Furthermore, random signals are mentioned; see Lecture II for relevant definitions.

Continuous time Wiener filters

The design of optimal filters in continuous time when random signals are involved can be found in numerous texts (e.g., reference [18]) under the title 'Wiener filtering'. With reference to the figure 16 where y and d are stationary random signals, we seek to estimate d by selection of filter $f(t)$ so that $E[e^2(t)]$ is minimised. If $f(t)$ is not restricted to being causal then the necessary condition that f should satisfy is the so-called Wiener-Hopf integral equation,

$$R_{dy}(\tau) = \int_{-\infty}^{\infty} R_{yy}(\tau - \alpha)f(\alpha)\,d\alpha \qquad \text{for all } \tau.$$

(An alternate statement is that the estimation error is uncorrelated with estimate $z(t)$.)

The acausal solution is obtained by Fourier transforming the equation to give

$$F(\omega) = \frac{S_{dy}(\omega)}{S_{yy}(\omega)} .$$

This is a general signal shaping formulation and it can be made into an inverse filtering problem by redrawing the figure as shown in fig. 17. Here measurement y is a noisy version of x and d is identified with x. The solution is (assuming n and x are uncorrelated)

$$F(\omega) = \frac{H(\omega)S_{xx}(\omega)}{S_{nn}(\omega) + |H(\omega)|^2 S_{xx}(\omega)}$$

Such a solution (being acausal) would require implementation using forward and reverse filtering. The imposition of the condition that the optimal system be *causal* results in a modified Wiener-Hopf equation, namely,

$$R_{dy}(\tau) = \int_0^\infty R_{yy}(\tau - \alpha)f(\alpha)d\alpha \qquad \text{for} \quad \tau > 0 \quad \text{only.}$$

The solution of this follows the method of spectral factorisation and is summarised below (see page 341 of reference [18]).

(i) Express $S_{yy}(-js) = A^+(s)A^-(s)$ where $A^+(s)$ is a minimum phase transfer function.

(ii) Form the ratio $S_{dy}(-js)/A^-(s) = B^+(s) + B^-(s)$ where $B^+(s)$ is analytic for $Re(s) > 0$.

(iii) The required $F(s) = B^+(s)/A^+(s)$ which is now causal.

Since the causality condition is an additional constraint, the value of the optimal error variance is greater for this case than the acausal case. Just as for the acausal case this can be phrased as an inverse filtering problem. Directly analogous results may be obtained in discrete time and are detailed with examples in reference [18], chapter 10, where the emphasis is on filtering noise from data rather than inverse filtering. We note finally in this section as we did earlier, that this optimisation procedure results in a filter that is a compromise between bandwidth and output signal-to-noise ratio.

Finite optimum (inverse) discrete time filters

The previous section related to the problem of calculating the optimal filter which in general will have infinite length impulse response. We shall now consider the problem of optimal *finite length* filters. As an introduction to this we consider the illustration below (see references [17] and [19] for further details).

Suppose $H(z) = (1 + kz^{-1})$ where $|k| < 1$ then $H^{-1}(z) = 1/1 + kz^{-1}$ and this inverse has an infinite impulse response. Let us now seek a truncated approximate inverse, but rather than simply truncating the (infinite) impulse response of the inverse we optimise the coefficients. For example, suppose a 'two coefficient inverse' is to be used, i.e., the impulse response sequence is (f_o, f_1), then the filter output is $(1, k) * (f_o, f_1) = (f_o, kf_o + f_1, kf_1)$, and the desired output is $(1, 0, 0)$ so the error sequence is $e^o = (1 - f_o, -kf_o + f_1, -kf_1)$. The error energy is

$$I = \sum_{n=0}^{2} e^2(n)$$

and this is minimised with respect to f_o and f_1, i.e., $\frac{\partial I}{\partial f_o}$; $\frac{\partial I}{\partial f_1} = 0$ to give the optimal coefficients. This procedure can be extended to higher order filters and leads to optimum inverse referred to as least error energy approximate inverses.

Having introduced the ideas we now consider the more general formulation depicted in fig. 18. The aim is to design a filter f, which, when applied to the signal y produces an output $d(n)$ which approximates a desired signal $d(n)$; Various formulations are possible including both deterministic and random sequences (see ref. [12]). For brevity we only mention the case of deterministic signals where the function to be minimised is the error energy $I = \Sigma e^2(n)$, where

$e(n) = d(n) - \sum_{k=0}^{N} f(k)y(n - k)$ and $f(k)$ is the filter impulse response function to be obtained. Inversion (of $h(n)$ say) is a special case of this obtained by setting $x(n) = d(n) = \delta(n)$ (see fig. 18). (an extension to this is to shift the delta function and write $x(n) = \delta(n)$ and $d(n) = \delta(n - n_o)$ (ref. [26]). This is one approach to overcoming the problem of inverting a non-minimum phase sequence.) Returning to the basic formulation a necessary condition for the minimisation of I is

$$\sum_{k=0}^{N} f(k) \sum_{n} y(n - k)y(n - i) = \sum_{n} d(n)y(n - i) \qquad i = 0, 1, \ldots, N$$

These equations are known as the *normal equations*. The specification of the range of summation of n leads to two distinct formulations (reference [12]) known as the autocorrelation and the covariance methods, respectively. There are several variations on this basic theme ultimately concerned with the specification of the optimal filter coefficients as the solution of a set of linear equations, often employing the so-called 'Levinson procedure' or the Durbin modification (references [20,21,24]).

In conclusion we note that other methods of deconvolution exist (e.g., reference [22]). We also particularly mention a technique referred to as 'homomorphic deconvolution' which is not discussed here since

it is considered in a later chapter on cepstral analysis.

REFERENCES/BIBLIOGRAPHY

A. General texts covering time series analysis ('Engineering' approach)

1. Applications of Time Series Analysis - Course Notes available from ISVR, University of Southampton.
2. K.G. Beauchamp. Signal Processing. George Allen and Unwin Ltd. 1973.
3. J.S. Bendat and A.G. Piersol. Random Data: Analysis and Measurement Procedures. Wiley Interscience. 1971.
4. R.K. Otnes and L. Enochson. Applied Time Series Analysis Vol. 1. Basic Techniques. Wiley and Sons. 1978.
5. R.B. Randall. Frequency Analysis. Bruel and Kjaer Publication. 1977.

(Statistical/mathematical approach)

6. G.M. Jenkins and D.G. Watts. Spectral Analysis and its Applications. Holden Day. 1968.

B. General texts covering digital signal processing

7. E.O. Brigham. The Fast Fourier Transform. Prentice Hall. 1974.
8. D. Childers and A. Durling. Digital Filtering and Signal Processing. West Publishing Co. 1975.
9. A.V. Oppenheim and R.W. Schafer. Digital Signal Processing. Prentice Hall. 1975.
10. L.R. Rabiner and B. Gold. Theory and Application of Digital Signal Processing. Prentice Hall. 1975.

C. Other references

11. F.J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. 1978 Proc. IEEE Vol. 66, No. 1.
12. J. Makhoul. Linear prediction: A tutorial review. 1975 Proc. IEEE Vol. 63, No. 4.
13. J.H. McClellan, T.W. Parks, L.R. Rabiner. A computer program for designing optimum FIR linear phase digital filters. 1973 IEEE Trans. Audio and Electroacoustics, Vol. AU-21, No. 6.
14. C.M. Rader. An improved algorithm for high speed autocorrelation with application to spectral estimation. 1970 IEEE Trans. Audio and Electroacoustics, Vol. AU-18.
15. F.R. Spitznogle and A.H. Quazi. Representation and analysis of time-limited signals using a complex exponential algorithm. 1970 Journal of the Acoustical Society of America Vol. 47, No. 5.
16. K. Steiglitz. Computer aided design of recursive digital filters. 1970 IEEE Trans. Audio and Electroacoustics Vol. AU-18.
17. E.A. Robinson. Statistical Communication and Detection. Griffin. 1967.
18. A. Papoulis. Signal Analysis. McGraw Hill Book Co. 1977.
19. S.M. Bozic. Digital and Kalman Filtering. Edward Arnold. 1979.
20. N. Levinson. The Wiener RMS error criterion in filter design and prediction. 1947 Journal Math. Physics. Vol. 25, No. 4.
21. J. Durbin. Efficient estimation of parameters in moving-average models. 1959 Biometrika, Vol. 46, Parts 1 and 2.
22. G. Thomas. Deconvolution and linear tracking problems. 1980 Signal Processing Vol. 2, No. 2.
23. J.D.E. Beynon and D.R. Lamb. Charge-Coupled Devices and their Applications. McGraw Hill. 1980.
24. J.F. Claerbout. Fundamentals of Geophysical Data Processing: With Applications to Petroleum Prospecting. McGraw Hill Book Co. 1976.
25. P.Y. Ktonas and N. Papp. Instantaneous envelope and phase extraction from real signals. 1980 Signal Processing 2.
26. S. Treitel and E.A. Robinson. The design of high resolution digital filters. IEEE Trans. on Geosci. Electronics, GE4. 1966.

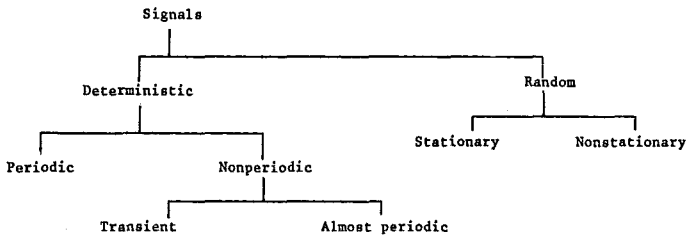


Fig. 1. Signal classification.

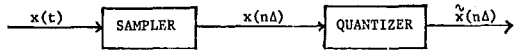


Fig. 2. Block diagram representation of analogue to digital conversion.

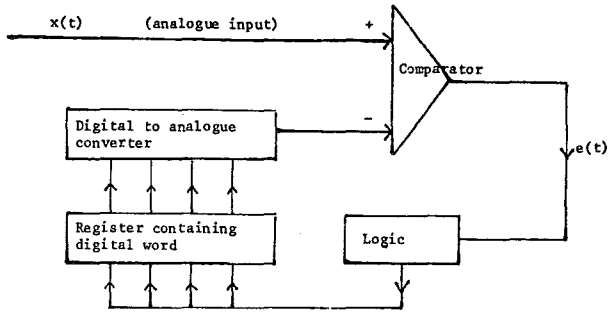


Fig. 3. An analogue to digital conversion scheme.

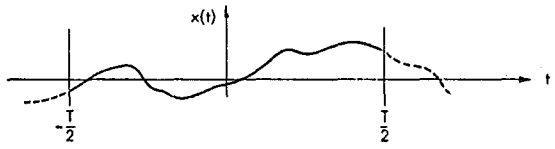


Fig. 4. A truncated (windowed) segment of data.

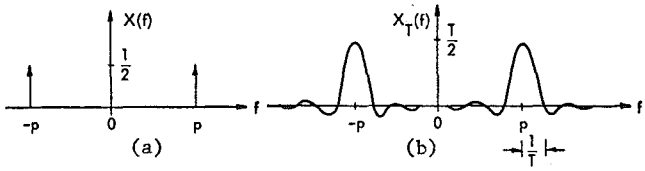


Fig. 5. Fourier transforms of a cosine wave (a) and a truncated cosine wave (b).

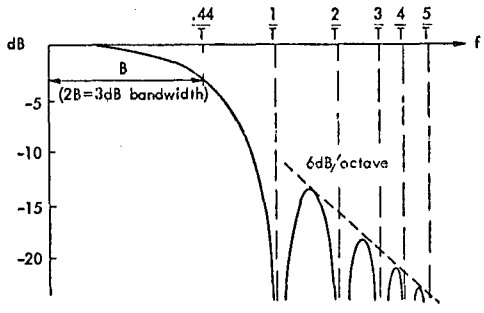


Fig. 6. Frequency domain characteristics of a rectangular data window.

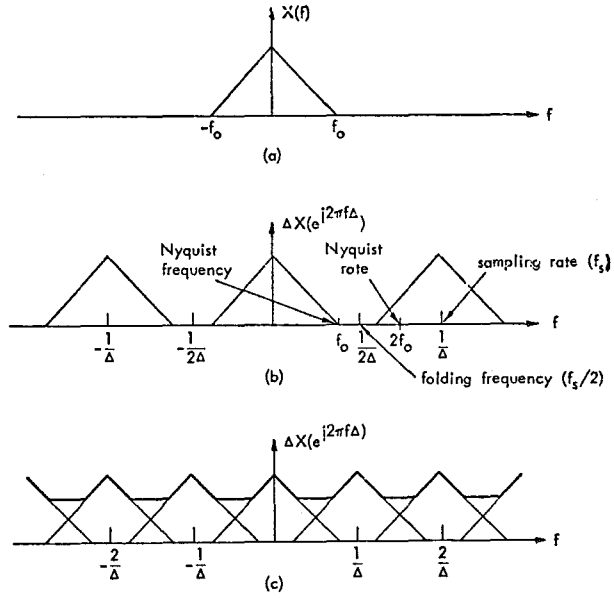


Fig. 7. The Fourier transform of a continuous wave (a) and the transform of corresponding sampled versions ((b) and (c)).

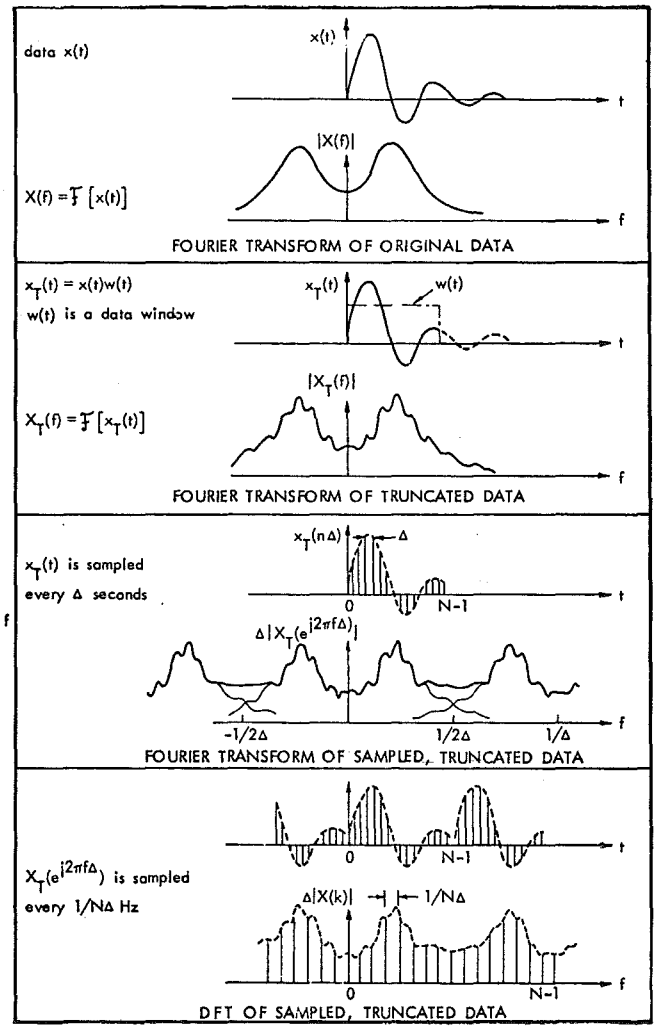


Fig. 8. Discrete Fourier transform of a truncated aperiodic function related (pictorially) to the transform X(f).

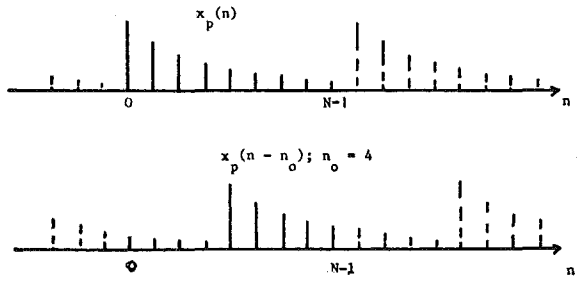


Fig. 9. A periodic sequence and a shifted version of the sequence.

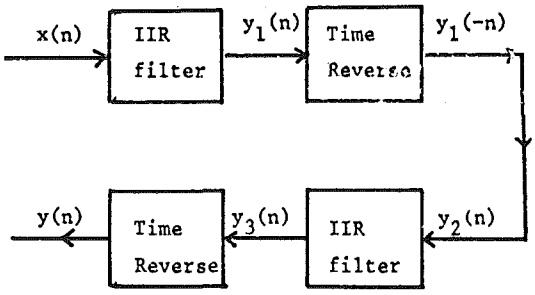


Fig. 10. Zero phase filtering using IIR filters.

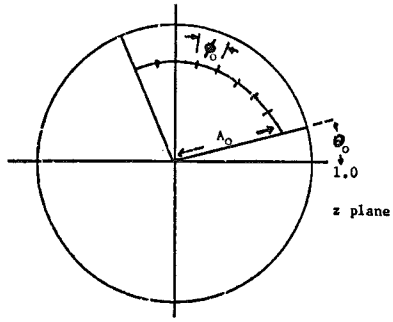


Fig. 11. Spiral contour in the z-plane for the CZT algorithm.

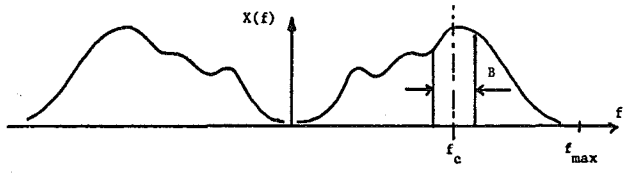


Fig. 12a. Bandwidth centred on frequency f_c requiring fine resolution.

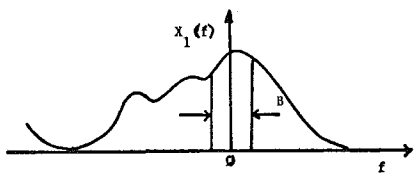


Fig. 12b. Fourier transform of 'heterodyned' signal.

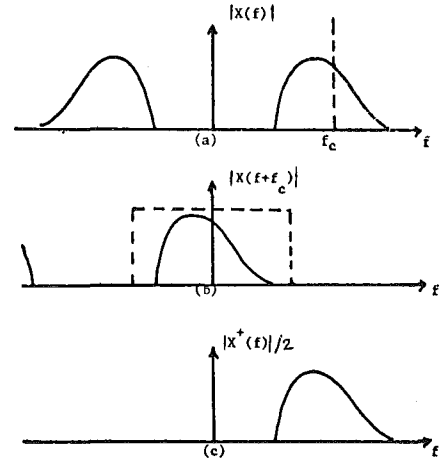


Fig. 13. Creation of the analytic signal x^+ from x.

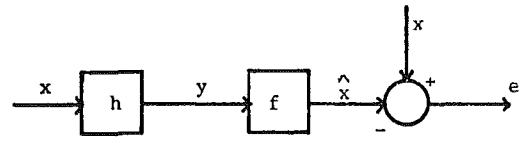


Fig. 14. Operator f is designed to operate on measurement of x to recover an estimate.

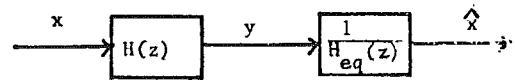


Fig. 15. Inversion of the $H_{eq}(z)$ portion of H(z) only.

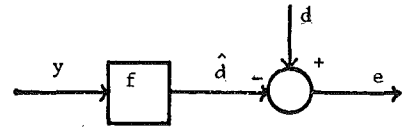


Fig. 16. Filter f operates on y to approximate d.

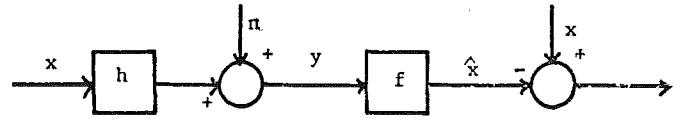


Fig. 17. Inverse filtering to recover x from noisy measurement y.

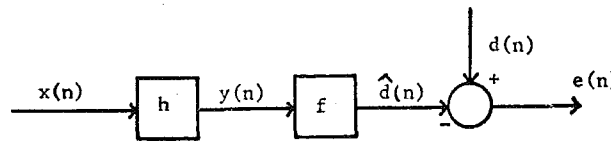


Fig. 18. Design of a finite optimum filter f.

LINEAR SIGNAL PROCESSING II

J.K. Hammond

Institute of Sound and Vibration Research
University of Southampton
Southampton SO9 5NH
England

SUMMARY

This is a companion chapter to the previous one. It is concerned with the treatment of random (stochastic) processes. This is often referred to as 'time series analysis' in the statistical literature. Whilst the computation methods referred to use digital techniques, the basic principles that will be outlined will refer to continuous time processes. Thus the concepts of periodicity, aliasing, etc. that are involved will not be mentioned explicitly.

Discussions will be restricted to stationary random processes and initially the definition of correlation functions and spectra are given. This is followed by a review of conventional estimation methods (implemented using digital schemes). Next the concepts of residual spectra and partial and multiple coherence functions for multi input systems are explained. Finally functions of more than one variable are treated. This relates directly to quantities that depend both on space and time. Aspects of two dimensional signal processing are reviewed and some applications are discussed.

1. STOCHASTIC PROCESSES

We begin with the description of nondeterministic signals. We may visualise a sample of such a signal as shown in fig. 1. The essential feature of such a signal is that its value at some future time cannot be predicted exactly. The elements of probability and statistics required to describe such processes are treated fully by Jenkins and Watts [4], see also refs. [1,2,3,17]. We shall only summarise some of the basic features here, using 'continuous time' for convenience. We start by defining a *sample space* (Ω), as the set of all possible outcomes of an *experiment* of chance. Any subset of points in the sample space is an *event* E, and individual elements (ω) in the sample space are called '*elementary events*'. A number is assigned to each event E in Ω that is a measure of the likelihood of occurrence of event E, called the *probability* of occurrence of event E written P(E). For any element ω in the sample space we can define a function $X(\omega)$ which is said to be a *random variable* defined on the sample space and both discrete and continuous random variable are described in terms of their probability distributions. It is now necessary to introduce 'time' into the picture. Consider a random variable which is time dependent. A possible notation to emphasise this dependence is $X_t(\omega)$, i.e., a 'parametered' random variable. This is abbreviated $X(t)$ and implies that for each time point t in some interval $X(t)$ exists having some defined sample space, and an associated probability density function $p_t(x)$ (or $p(x(t)$ or $p(x, t)$). Such a parametered random variable is called a *stochastic process* or *time series* and is sometimes denoted $\{X(t)\}$. The concept of the '*ensemble*' is often introduced. An observed time history may be regarded as one realisation of an infinite number of records, i.e., records which may be visualised as being generated by a large number of independent repetitions of the same experiment under conditions which are essentially identical.

The *probability density function* for the stochastic process is $p(x, t) = \lim_{\delta x \to 0} \frac{\text{Prob}[x < X(t) \leq x + \delta x]}{\delta x} = \frac{dF(x,t)}{dx}$ where $F(x, t) = \text{Prob}[X(t) \leq x]$ is the *cumulative distribution function*. Similarly, a joint probability density function $p(x_1, t_1; x_2, t_2)$ is defined as $\partial^2 F(x_1, t_1; x_2, t_2)/\partial x_1 \partial x_2$ where $F(x_1, t_1; x_2, t_2)$ is the $\text{Prob}[X(t_1) \leq x_1, X(t_2) \leq x_2]$. For a bivariate stochastic process $\{X(t), Y(t)\}$ the density function is $p(x, t_1; y, t_2)$.

Moments of the stochastic process are defined in the usual way for random variables with the time dependence the only difference, e.g., the mean is defined

$$\mu_x(t) = E[X(t)] = \int_{-\infty}^{\infty} x\, p(x, t)\, dx$$

and variance is

$$\sigma_x^2(t) = E[(X - \mu_x(t))^2] = \int_{-\infty}^{\infty} (x - \mu_x(t))^2 p(x, t)\, dx.$$

$E[\ ]$ denotes the expected value operator which may be identified with averaging *across* the set of records or 'ensemble averaging'.

Stationarity

In general the properties of a stochastic process are time dependent but to simplify matters we often assume that a sort of 'steady state' has been reached in the sense that the statistical properties are unchanged under a shift in time, i.e., (i) $p(x, t) = p(x)$ (implying μ_x and σ_x are constant); (ii) $p(x_1, t_1; x_2, t_2)$ is a function of $t_2 - t_1$ only and not both t_1 and t_2. If only conditions (i) and (ii) hold, the process is said to be 'weakly stationary' or simply stationary. If similar conditions for all higher order joint probabilities hold, then the process is completely stationary. We shall be content with conditions (i) and (ii). The assumption of stationarity (though an approximation) has important consequences, since it is the key to being able to replace our previous 'ensemble' averages across a *set* of records with 'time averages' along a single record.

1.1 Covariance (Correlation) Functions

A widely used average for a random process $\{X(t)\}$ is the *autocovariance function* defined as $R_{xx}(t_1, t_2) = E\{(X(t_1) - \mu_x(t_1))(X(t_2) - \mu_x(t_2))\}$. This is a measure of the degree of association of the signal at time t_1 with itself at time t_2. In the engineering literature the quantity

$E[X(t_1)X(t_2)]$ is usually referred to as the autocorrelation function. If a process is stationary, we note that the statistical properties are unchanged under a time shift so that $R_{xx}(t_1, t_2) = R_{xx}(t_2 - t_1)$ or in more usual notation $R_{xx}(\tau)$ (τ is the 'lag'). For a bivariate process the *cross covariance* (cross correlation) function is a measure of the degree of association of signal X at time t_1 with signal Y at t_2. We will assume zero mean processes for simplicity of notation and so $R_{xy}(t_1, t_2) = E[X(t_1)Y(t_2)]$ and if the processes are stationary $R_{xy}(\tau) = E[X(t)Y(t + \tau)]$. The averaging operations (so far) should be interpreted as ensemble averages.

Time averages

The definitions above have all involved the underlying probability distributions (or equivalently ensemble averaging). Often only a single record is available from which to make estimates of the averages we have defined. This raises the question as to whether averages *along* a record, i.e., a time average, might be used in place of the ensemble average. It seems self evident that 'stationarity' is a prerequisite if this is to be feasible, and it turns out that for stationary processes *some* time averages approach the equivalent ensemble average as record lengths increase. Such averages are said to be *ergodic*, and they should be correctly discussed via estimation theory. However, we shall anticipate that stationary processes are ergodic with respect to the mean and auto and cross covariance/correlation functions and note that this has prompted many texts to adopt the following 'definitions'.

(i) The mean value $$\mu_x = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, dt; \tag{1}$$

(ii) The cross correlation function $$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)y(t + \tau)dt \tag{2}$$

with $y = x$ for the autocorrelation function.

1.2 Spectral Density Functions

Consider the sample function of a random process $x_T(t)$ as in figure 1 then if $X_T(f)$ denotes the Fourier transform of $x_T(t)$ the average power is

$$\frac{1}{T}\int_{-\infty}^{\infty} x_T^2(t)dt = \frac{1}{T}\int_{-\infty}^{\infty} |X_T(f)|^2 df.$$

The quantity $|X_T(f)|^2/T$ is called the sample (raw) spectral density function denoted $\hat{S}_{xx}(f)$ Simply defining $\lim_{T\to\infty} (|X_T(f)|^2/T)$ as the power spectral density of the process is useless since the quantity does not converge in a statistical sense, i.e., the values of this are just as erratic for long data lengths as shorter ones. The use of 'more data' does not 'improve' the estimate (indeed, the estimate of what?). We have an estimate for which the ergodic property does not hold. However, we can introduce a useful quantity by averaging this to remove the erratic behaviour and define

$$S_{xx}(f) = \lim_{T\to\infty} \frac{E|X_T(f)|^2}{T} \tag{3}$$

so (from the equation above)

$$\sigma_x^2 = \int_{-\infty}^{\infty} S_{xx}(f)df. \tag{4}$$

$S_{xx}(f)$ is the power spectral density function and describes the decomposition of the average power of $X(t)$ over frequency.

It follows from the definition (3) that $S_{xx}(f)$ is related to $R_{xx}(\tau)$ by

$$S_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau)e^{-j2\pi f\tau}d\tau \qquad \text{and} \qquad R_{xx}(\tau) = \int_{-\infty}^{\infty} S_{xx}(f)e^{j2\pi f\tau}df. \tag{5}$$

The *cross spectral density* is similarly defined as

$$S_{xy}(f) = \lim_{T\to\infty} E\left\{\frac{X_T^*(f)Y_T(f)}{T}\right\} \tag{6}$$

and is related to the cross correlation function by

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau)e^{-j2\pi f\tau}d\tau \qquad \text{and} \qquad R_{xy}(\tau) = \int_{-\infty}^{\infty} S_{xy}(f)e^{j2\pi f\tau}df. \tag{7}$$

The cross spectral density is complex and may be expressed $|S_{xy}(f)|e^{j \arg S_{xy}(f)}$. $|S_{xy}(f)|$ is the cross amplitude spectral density and indicates whether frequency components in one time series are associated with large or small amplitudes at the same frequency in the other series. Arg $S_{xy}(f)$ is the phase spectral density and indicates whether frequency components in one series lag or lead the components at the same frequency in the other series.

1.3 Input-output Relationships for Linear Systems

A linear, time invariant system may be characterised by its impulse response function $h(t)$ and if $X(t)$ is the input and $Y(t)$ the output to an input starting at $t = t_o$ then

$$Y(t) = \int_{t_o}^{t} h(t - t_1)X(t_1)dt_1. \tag{8}$$

If $X(t)$ is stationary, t_o approaches $-\infty$ and the system is stable then $Y(t)$ is stationary and

$$R_{yy}(\tau) = \int_{o}^{\infty}\int_{o}^{\infty} h(\tau_1)h(\tau_2)R_{xx}(\tau + \tau_1 - \tau_2)d\tau_1 d\tau_2 \tag{9}$$

which relates input and output correlation functions. Fourier transforming this equation results in the simpler spectral density input output equation

$$S_{yy}(f) = |H(f)|^2 S_{xx}(f) \tag{10}$$

where $H(f)$ is the system frequency response function. This equation indicates how the input spectral density is 'shaped' by $|H(f)|^2$.

'Cross' relations involving input and output can be obtained similarly. The time domain form is

$$R_{xy}(\tau) = \int_{o}^{\infty} h(\tau_1)R_{xx}(\tau - \tau_1)d\tau_1 \tag{11}$$

and the frequency domain equivalent is

$$S_{xy}(f) = H(f)S_{xx}(f) \tag{12}$$

Since phase information is retained in these relationships they are sometimes used in system identification schemes where only input and output are measureable and the system is unknown. In particular if the input is white noise then the cross correlation function is proportional to the system impulse response.

The ordinary coherence function

The ordinary coherence function between input $X(t)$ and output $Y(t)$ is defined as

$$\gamma_{xy}^2(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)\, S_{yy}(f)} \tag{13}$$

and can be shown to satisfy $0 \leqslant \gamma_{xy}^2(f) \leqslant 1$. If X and Y are completely unrelated, $\gamma_{xy}^2(f) = 0$, whilst if X and Y are linearly related then $\gamma_{xy}^2(f) = 1$. In fact the coherence function arises naturally when given two random processes X and Y, one asks the question 'Is it possible to produce Y (in some approximate sense) from purely linear operations on X?' This is a problem of system identification. Put another way, we wish to find the impulse response function $h(t)$ such that the variance of the 'error' is minimised (see fig. 2). Such an 'optimal filter' satisfies the Wiener-Hopf equation (written here for the case where $h(t)$ need not be causal)

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} h(\tau_1)R_{xx}(\tau - \tau_1)d\tau_1. \tag{14}$$

From this it can be argued that the spectral density of the error is $S_{ee}(f) = S_{yy}(f)(1 - \gamma_{xy}^2(f))$ and $\gamma_{xy}^2(f)$ can be written $\gamma_{xy}^2(f) = S_{zz}(y)/S_{yy}(f)$ showing that the coherence function is a measure of the proportion of the power of Y that can be accounted for by linear operations on the input. (This is taken further in section 3.)

Discrete time processes

The treatment above has been restricted exclusively to continuous time processes but the development could equally well have been done in discrete time. For example, the cross correlation of two sequences might be defined

$$R_{xy}(m) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)y(n + m)$$

and the cross spectral density defined as

$$S_{xy}(f) = \sum_{m=-\infty}^{\infty} R_{xy}(m)e^{-j2\pi mf}$$

and this can be inverted as

$$R_{xy}(m) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{xy}(f)e^{j2\pi fm}df.$$

It is important to note that $S_{xy}(f)$ is periodic in f with period unity (since we interpret the samples as occurring every second in this particular case). Input-output relationships, stateforms (see 3.1) etc. may be redeveloped using a completely discrete formulation.

2. ESTIMATION METHODS AND ESTIMATION ERRORS

In this section we introduce and discuss estimation methods for random signals based on a single recording of the process.

2.1 Estimator Errors and Accuracy

Suppose the mean value of a random process is μ_x (unknown to us) and we only have a record length T of the signal. A reasonable estimate $\bar{x}$ of μ_x is

$$\bar{x} = \frac{1}{T}\int_0^T x(t)dt.$$

The value obtained is a sample value of a random variable $\bar{X}$, having its own probability distribution (called a sample distribution) and $\bar{x}$ is a single realisation. Each $\bar{x}$ computed from a (different) record of length T will, in general, be different. If the estimation procedure is to be useful we would hope that (i) the scatter of values of $\bar{x}$ is not 'too great' and should lie 'close' to μ_x; (ii) the more data we take (the larger T is) the 'better' the estimate. We can formalise these ideas as follows (see refs. [2, 4]).

Let ϕ be a parameter we wish to estimate (in the introductory example above ϕ is μ_x) and let $\hat{\Phi}$ be an estimator for ϕ, then $\hat{\Phi}$ is a random variable with its own sampling distribution $p(\hat{\phi})$ which is often difficult to obtain and so we shall settle for a few summarising properties.

(1) *Bias error*. The bias of an estimator is defined $b(\hat{\Phi}) = E[\hat{\Phi}] - \phi$. This is the difference between the average of the estimator and the true value. If $b(\hat{\Phi}) = 0$, then $\hat{\Phi}$ is 'unbiased'. Whilst it might seem desirable to use an unbiased estimator it is sometimes prudent to tolerate some bias if it means that the variability of the estimate is reduced (relative to that of an unbiased estimator).

(2) *Variance*. The variance of an estimator is a measure of the spread of $\hat{\Phi}$ about its *own* mean value $\text{Variance}[\hat{\Phi}] = E[(\hat{\Phi} - E(\hat{\Phi}))^2]$. The square root of the variance is the standard deviation $(\sigma(\hat{\Phi}))$ of the estimator. We would hope that this would be 'small', i.e., the probability density function should be 'peaky', but this requirement often increases the bias. A measure of the relative importance of bias and variance is the mean square error.

(3) *Mean square error*. (mse). The mse of an estimator is a measure of the spread of values of $\hat{\Phi}$ about the value ϕ, i.e., $\text{mse}[\hat{\Phi}] = E[(\hat{\Phi} - \phi)^2]$. This can be rewritten as $\text{mse}[\hat{\Phi}] = \text{Variance}[\hat{\Phi}] + b^2[\hat{\Phi}]$ showing that mse reflects both variance and bias. If the mse reduces as the sample size (amount of data used) increases then the estimator is *consistent*.

Confidence intervals

The estimates $\hat{\phi}$ referred to so far are *point estimates*, i.e., single values. It is frequently desirable to know that a parameter is likely to fall within a certain *interval of values*. For example, if we estimate a mean value $\bar{x}$ as 5, perhaps it is 'likely' that μ_x lies in the interval 4.5-5.5. These estimates are *interval estimates*, and are called *confidence intervals* when we attach a number describing the likelihood of the parameter falling within the interval (see section 2.2).

2.2 Estimators for Stochastic Processes

The problem of the study of estimators for stochastic processes is a wide one and some idea of the extent of the field can be obtained from [4]. Several other texts are not as detailed and give useful summaries; these are [1,2,3,17]. We shall pick out the autocorrelation function and power spectral density as examples, and try and indicate some of the features that should be considered when analysing a sample time history. Following this we briefly discuss cross spectra and system identification. Unless otherwise stated we shall assume we are dealing with realisations of a continuous, stationary random process. The data is sampled so that the computation may be done and is therefore subject to the attendant 'aliasing' problems. It is assumed throughout that the sampling rates are 'sufficiently high'.

The Autocorrelation Function

(a) *Definition*. If we have a signal defined for $0 \leq t \leq T$, then there are two (ref. [4]) commonly used estimators of the function $R_{xx}(\tau)$. Following [4] we write tham as $\hat{R}_{XX}(\tau)$ and $\hat{R}'_{XX}(\tau)$ where (assuming a zero mean process)

$$\hat{R}_{XX}(\tau) = \frac{1}{T}\int_0^{T-|\tau|} X(t)X(t+|\tau|)dt \qquad 0 \leq |\tau| \leq T, \quad \text{and} \quad = 0 \qquad |\tau| > T \tag{15}$$

and

$$\hat{R}'_X(\tau) = \frac{1}{T-|\tau|}\int_0^{T-|\tau|} X(t)X(t+|\tau|)dt \qquad 0 \leq |\tau| \leq T, \quad \text{and} \quad = 0 \qquad |\tau| > T. \tag{16}$$

The 'engineering literature' seems to favour the latter since it is unbiased but it is the mse that should be the criterion for a choice between the two.

(b) *Statistical considerations*. For τ in the range $0 \leq |\tau| \leq T$, $\hat{R}'_{XX}$ is unbiased, and $E[\hat{R}_{XX}(\tau)] = R_{xx}(\tau)\left[1 - \frac{|\tau|}{T}\right]$ showing (15) to be 'asymptotically unbiased'. An approximate expression for the variance of $\hat{R}_{XX}(\tau)$ is

$$\text{Var}[\hat{R}_{XX}(\tau)] \simeq \frac{1}{T}\int_{-\infty}^{\infty} [R^2_{xx}(r) + R_{xx}(r+\tau)R_{xx}(r-\tau)]dr$$

showing this to be a consistent estimator. When $\tau << T$ there is little to choose between the two estimators, but as τ approaches T the variance of the unbiased estimator diverges. Another important feature of the estimators is that values computed for neighbouring values of τ are sample values of (in general) strongly correlated random variables meaning that $\hat{R}_{XX}(\tau)$ is more strongly correlated than the original time series $x(t)$ so the estimate may not decay as rapidly as one might expect (see section 5.3 of ref. [4]).

(c) *Computation.* Computation of (15) is achieved using sampled data with the formula

$$\hat{R}_{xx}(m) = \frac{1}{N}\sum_{n=0}^{N-|m|-1} x(n)x(n + |m|). \qquad (17)$$

The divisor is $N - |m|$ for $\hat{R}'_{XX}(m)$; m is the lag and may take values $0 \leq m \leq M-1$. The calculations to be done in (17) are time consuming if M is not small and FFT methods can be employed with considerable speed advantage. A method due to Rader [7] is particularly effective, the procedure being based on the idea of 'fast' convolution.

The power spectral density

Estimation methods for spectra considered here will relate to the 'traditional' methods rather than the recently developed 'parametric' methods. The standard methods are (i) the Fourier transform of the autocorrelation function (indirect method); (ii) the direct method of averaging periodograms; (iii) filtering (either bandpass or complex demodulation (heterodyning) and low pass filtering). We shall outline (i) and (ii). (An interesting and readable paper [5] considers various aspects (computational and statistical) of spectral estimation though recently [10] some criticisms of the tendency to favour method (ii) have been voiced.)

(i) *The Fourier transform of the autocorrelation function*

(a) *Definition.* An estimate of the power spectral density from a length of data T (here assumed to be defined for $|t| \leq T/2$) is

$$\hat{S}_{XX}(f) = \frac{|X_T(f)|^2}{T} = \int_{-T}^{T} \hat{R}_{XX}(\tau)e^{-j2\pi f\tau}d\tau. \qquad (18)$$

This is termed the 'raw' spectral density since it turns out (see (b) which follows) that the variability of this estimator is independent of the data length T and so is no use as an estimator. To reduce the sampling fluctuations the estimate must be smoothed in the frequency domain and this can be achieved by multiplying the autocorrelation function estimate by a 'lag window' $w(\tau)$ which has a Fourier transform $W(f)$ (a 'spectral window'), and Fourier transforming the product. This defines the estimator $\tilde{S}_{XX}(f)$ where

$$\tilde{S}_{XX}(f) = \int_{-T}^{T} \hat{R}_{XX}(\tau)w(\tau)e^{-j2\pi f\tau}d\tau \qquad (19)$$

which is the convolution of the 'raw' spectral density with $W(f)$, i.e., $\hat{S}_{XX}(f)*W(f)$. This estimation procedure may be viewed as a smoothing operation in the frequency domain or, in the time domain, the lag window can be regarded as reducing the 'importance' of values of $\hat{R}_{XX}$ as τ increases (since the variability of the estimator $\hat{R}_{XX}$ is increased as τ increases).

(b) *Statistical properties.* (See references [3,4,17] for details).

Bias. The raw spectral density $\hat{S}_{XX}(f)$ is an asymptotically unbiased estimator.

Variance. Since $\hat{S}_{XX}(f)$ is a squared quantity (and if we assume $X(t)$ is Gaussian) then it can be argued that $2\hat{S}_{XX}(f)/S_{XX}(f)$ is distributed as a chi-squared random variable with two degrees of freedom, i.e., χ^2_2 (for all values of data length T). It follows that the variance of $\hat{S}_{XX}(f)$ is equal to $S^2_{XX}(f)$. So $\hat{S}_{XX}(f)$ is an inconsistent estimator of $S_{XX}(f)$, and the standard deviation of the estimate is as large as the quantity being estimated. It is these unsatisfactory properties that lead to the consideration of another estimate.

Smoothed spectral estimators. Since $\tilde{S}_{XX}(f)$ is a convolution of $\hat{S}_{XX}(f)$ with $W(f)$ it turns out that $\tilde{S}_{XX}(f)$ is biased. In fact (for large T)

$$E[\tilde{S}_{XX}(f)] \simeq \int_{-\infty}^{\infty} S_{xx}(g)W(f - g)dg \qquad \text{and} \qquad \text{bias}[\tilde{S}_{XX}(f)] \simeq \int_{-\infty}^{\infty} [w(\tau) - 1]R_{xx}(\tau)e^{-j2\pi f\tau}d\tau. \qquad (20)$$

The first expression indicates how the average of the estimate is distorted by smearing and 'leakage' of the spectral density owing to the spectral window width and shape (side lobes). The general effect is to reduce the dynamic range of the spectra, i.e., peaks are underestimated and troughs are overestimated and this effect is reduced as the spectral window gets narrower. However, the windows cannot be made too narrow since then there is little smoothing and the random errors increase. Consequently there is always the problem of trading bias errors against random errors. (The bias problem is sometimes referred to under 'bandwidth considerations').

Since $\tilde{S}_{XX}$ is a weighted sum of values of $\hat{S}_{XX}$ is is argued [4] that $n\tilde{S}_{XX}(f)/S_{XX}(f)$ is approximately distributed as χ^2_n where n (the number of degrees of freedom) $= 2BT$ and B is the resolution bandwidth. It follows that

$$\text{Var}[\tilde{S}_{XX}(f)] = \frac{S^2_{xx}(f)}{n/2} = \frac{S^2_{xx}(f)}{BT},$$

and in chapter 6 of ref. [4] details of estimator accuracy are tabulated for various lag windows.

The choice of window should depend on whether the concern is for statistical *stability* (low variance) or small bias. If the bias is small for all f, then we have high *fidelity*. Of course in a general situation we must be careful with both. For example, if the spectral function has narrow peaks of importance we may willingly tolerate some loss of stability to resolve the peaks properly, whilst if the spectral function is smooth then bias errors are not likely to be so important.

Confidence intervals. We now give 'interval estimates' based on the 'point estimates' for the smoothed spectral density. Since $n\hat{S}_{xx}(f)/S_{xx}(f)$ is distributed as a χ^2_n random variable then for a particular point estimate $\hat{S}_{xx}$ the $100(1-\alpha)\%$ confidence limits for S_{xx} are

$$\frac{n}{x_{n,\,1-\alpha/2}}\hat{S}_{xx}, \qquad \frac{n}{x_{n,\,\alpha/2}}\hat{S}_{xx},$$

and the confidence *interval* is the difference between them. Note that on a linear scale the confidence interval depends on the estimate, but on a log scale the inverval is $\log(n/x_{n,\alpha/2}) - \log(n/x_{n,1-\alpha/2})$ which is a constant independent of $\hat{S}_{xx}$.

(c) *Computation* of $\hat{S}_{xx}(f)$ is achieved by implementing a discrete form of (19). Reference [3] contains details.

(ii) *Segment averaging*

A method of spectral estimation that has become very popular (owing mainly to its speed of computation) is that of segment averaging discussed in Jenkins and Watts' book (section 6.3.4) as Bartlett's procedure and in some detail by Welch [8] The books by Bendat and Piersol [2], Otnes and Enochson [3] and Beauchamp [1] also discuss it in detail and emphasise it as the current main method of estimating spectra. The basic procedure is outlined with reference to figure 3.

A data length T is split up into q segments of length T_r (in this case non overlapping) and the raw periodogram formed for each segment, i.e., $\hat{S}_{xx_i}(f) = 1/T_r^r|X_{T_{ri}}(f)|^2$ for i = 1, ..., q, and to reduce the fluctuations the average

$$\hat{S}_{xx}(f) = \frac{1}{q}\sum_{i=1}^{q}\hat{S}_{xx_i}(f)$$

is formed and it can be argued (very approximately) that since each $2\hat{S}_{XX_i}/S_{xx}$ is a χ^2_2 random variable then $2\hat{S}_{XX}q/S_{xx}$ is χ^2_{2q} from which it follows that

$$\mathrm{Var}[\hat{S}_{XX}] \simeq \frac{S^2_{xx}(f)}{q}$$

and since the resolution bandwidth B is approximately $1/T_r$ then $\mathrm{Var}[\hat{S}_{XX}] \simeq \frac{S^2_{xx}(f)}{BT}$ (as before). Whilst this summarises the essential features of the method, references [8] and [5] give more elaborate procedures and insight. An important aspect is the use of data windows (or linear tapering as it is sometimes called) that are assigned to each segment. Furthermore, the segments may be chosen to overlap each other. The criticism of this method [10] is that linear windowing in effect ignores some data because of the tapered shape of the window. Intuitively the overlapping of segments compensates for this in some way and Welch [8] has results for this case though it is not easy to relate these results to the indirect method of Fourier transforming the autocorrelation function.

The use of a data window on a segment before Fourier transformation reduces leakage but has the disadvantage of also reducing the total energy of that segment and this must be compensated. This results in the calculation of modified periodograms for each segment of the form

$$\hat{S}_{xx_i}(f) = \frac{\frac{1}{T_r}\left|\int_{i\text{'th interval}} x(t)w(t)e^{-j2\pi ft}dt\right|^2}{\frac{1}{T_r}\int_{-T_r/2}^{T_r/2} w^2(t)dt}$$

where the denominator compensates for the 'energy reduction'. As far as computation is concerned the DFT (FFT) is used to compute the required modified periodograms and the appropriate summations performed.

Cross Spectra, Coherence and Transfer Functions

(a) *Cross spectra*

The basic considerations above also relate to cross spectral estimation together with some additional features. Detailed results can be found in chapter 9, ref. [4]. If the smoothed cross spectral density $\hat{S}_{xy}(f)$ is written as $|\hat{S}_{xy}(f)|e^{j\arg \hat{S}_{xy}(f)}$ then the variances of the amplitude ($|\hat{S}_{xy}|$) and phase ($\arg \hat{S}_{xy}$) estimators are proportional to 1/BT but are also strongly dependent on the coherency $\gamma^2_{xy}(f)$.

(b) *The coherence function*

The estimate for the coherence function is made up from estimates of smoothed power and cross spectra, i.e.,

$$\hat{\gamma}^2_{xy} = \frac{|\hat{S}_{xy}(f)|^2}{\hat{S}_{xx}(f)\hat{S}_{yy}(f)}.$$

(It should be noted that if 'raw' spectra are used in this estimate it is a simple matter to verify that the sample coherence function is unity for all frequencies for any signals x and y.) Statistical errors and confidence limits are given in [1,2,4] and a recent paper [6] discusses the bias of the

estimator. Results derived by Jenkins and Watts show that the bias is proportional to the square of the derivative of the phase spectrum between x and y. This means that the estimator is sensitive to delays between x and y and can also be expected to be inaccurate when, for example, x and y are input and output for a lightly damped oscillator. The bias error can be reduced by improving the resolution of the estimate. In fig. 4 the coherence function is estimated for simulated input/output results for an oscillator with centre frequency f_o = 1 Hz. The theoretical value of coherence is unity and the resolutions used are shown on the figure.

(c) *Frequency response functions* are estimated using smoothed spectra from the expression $\hat{H}(f) = \hat{S}_{xy}(f)/\hat{S}_{xx}(f)$, and results for errors and confidence limits can be found in [2,4]. The use of this nonparametric approach for systems with feedback is treated in [9]. (If feedback loops are not correctly accounted for the estimate of the frequency response function of an element in the loop can be biased.)

3. MULTIPLE INPUT SYSTEMS, RESIDUAL SPECTRA AND PARTIAL COHERENCE

Multi-input multi-output systems may be treated in the same manner as single-input single-output systems in section 1 and the results look similar as long as the appropriate vector-matrix interpretation is adopted. If a system characterised by the matrix of impulse response functions $H_I(t)$ has m inputs (vector $\underline{g}(t)$ and n outputs (vector $\underline{x}(t)$), then

$$\underline{x}(t) = \int_0^\infty H_I(\tau)\underline{g}(t-\tau)d\tau \qquad \text{(N.B. We shall use lower case letters for functions of time hereafter.)}$$

resulting in the spectral input-output relationship $S_{xx}(f) = H_F^*(f)S_{gg}(f)H_F^T(f)$. $S_{xx}(f)$, $S_{gg}(f)$ are the output and input spectral density matrices respectively (which are Hermitian) and $H_F(f)$ denotes the matrix of system frequency response functions. The input-output cross spectral relationships are $S_{gx}(f) = S_{gg}(f)H_F^T(f)$. These equations above are easily derived but the results themselves require further consideration if we are to use them to reveal details about the system; for example it may be of interest to know the relative importance of the inputs with reference to one particular output or indeed whether the set of inputs considered is 'adequate'. These and other problems associated with systems with more than one input will be discussed in section 3.2.

3.1 State Forms for Linear Systems

A widely used method of analysis for stochastically driven dynamic systems uses the state variable description (ref. [11]). This has achieved considerable attention in the treatment of control systems. It starts from the premise that a large class of processes can be modelled by the first order vector matrix equation

$$\dot{\underline{x}} = A\underline{x} + B\underline{w}; \qquad \underline{x}(t_o) \quad \text{and} \quad \underline{y} = C\underline{x}. \tag{21}$$

$\underline{x}$ is an n-dimensional state vector and $\underline{w}$ an m-dimensional vector of white noises with $E[\underline{w}(t_1)\underline{w}^T(t_2)] = Q\delta(t_1 - t_2)$ say, and $\underline{y}$ is a p-dimensional output vector. A, B, C are matrices of appropriate dimension[2] (which may be time varying).

The solution to this equation for state x(t) is

$$\underline{x}(t) = \Phi(t, t_o)\underline{x}(t_o) + \int_{t_o}^{t} \Phi(t, t_1)B\underline{w}(t_1)dt_1 \tag{22}$$

where Φ is the state transition matrix for the system. If $P(t) = E[\underline{x}(t)\underline{x}^T(t)]$ denotes the zero lag covariance matrix for the state then it is easy to verify that

$$\dot{P} = AP + PA^T + BQB^T; \qquad P(t_o) \tag{23}$$

which is true even if A, B are time dependent, showing that P may be computed by solving a linear differential equation directly rather than needing Φ. This method therefore offers a particularly simple approach to certain nonstationary situations, and will be discussed in a later lecture. The apparent restriction that the input $\underline{w}$ should be 'white' can be circumvented as long as whatever input there is can itself be regarded as the output of a 'white excited' linear system, i.e., a solution of the 'shaping filter' problem. This automatically means however that the dimension of the system is increased by the dimension of the shaping filter. In fact, retaining the 'white' excitation ensures that the state is a 'Markov vector' which in turn has important consequences insofar as the probabilistic structure of the state is concerned. For example, it is then possible to show that the transitional probability density function satisfies a partial differential equation called the Fokker-Planck equation which offers one approach to the treatment of nonlinear systems.

3.2 Residual Spectra

Recently interest has been shown in the application of the concepts of residual spectra and partial and multiple coherence for systems with more than one input. The objective of this section is to explain the significance and use of some of these concepts and draws extensively on the work reported by Dodds and Robson [16] and Bendat [12-15] and the recent book [17]. It is necessary to recall the concept of coherence as discussed in section 1.3. Suppose two measurements, x and y are available and one wishes to establish a linear transfer characteristic linking the two that account for y(t). Let $y_o(t)$ denote a stationary random process produced by operating on x(t) linearly, i.e.,

$$y_o(t) = \int_0^\infty h(\tau)x(t-\tau)d\tau$$

and $e(t) = y(t) - y_o(t)$ denote the error. The problem of system identification now becomes that of

finding the transfer function $h(t)$ that minimises $E[y^2(t)]$. This problem can be solved in the time domain using variational methods and results in the equation $E[e(t)x(t')] = 0$ for $t' < t$, or, in words, the estimation error is uncorrelated with past observations (the condition $t' < t$ results from the imposed causality condition on $h(t)$). This is called the Wiener-Hopf equation and is often written

$$R_{xy}(\tau_1) = \int_0^\infty h(\tau_2)R_{xx}(\tau_1 - \tau_2)d\tau_2 \qquad \tau_1 > 0$$

and the solution of this equation for $h(\tau)$ yields the required transfer function. Owing to the added condition $\tau_1 > 0$ the solution cannot be accomplished simply by Fourier transforming the equation, but requires the method of spectral factorisation. If, however, the requirement of causality is relaxed, then Fourier transforming the equation yields $H(f) = S_{xy}(f)/S_{xx}(f)$. Following Dodds and Robson [16] one may regard $y(t)$ as being composed of two parts y_2 and y_3 (see fig. 5). y_2 is fully coherent with y_1 and y_3 is uncoherent with y_1. Furthermore it is easily verified that $S_{y_2y_2}(f) = \gamma^2_{xy}S_{yy}(f)$ and $S_{y_3y_3}(f) = (1 - \gamma^2_{xy})S_{yy}(f)$.

It is useful to think of y_3 as a residual random variable, i.e., the result of y (or x_2) having had the (linear) effects of x_1 removed in an optimal (least squares) way, and so one might write $y_3 = x_{2.1}$ for this residual variable.

We extend the discussion to a two input-single output system. A simple system of this nature allows the method to be explained easily; the equations developed can be extended appropriately to describe systems with more than two inputs.

A two-input single-output system

Suppose three random processes x_1, x_2 and x_3 are measured (where one might choose to interpret one of these as an 'output' y, for example, x_3) and suppose the objective is to identify some form of excitation-response relationship between them. As noted in [2], when a single component of a random process is influenced by (many) others a distinct response relationship connecting it with any one of them is likely to be obscured by the action of the remainder. It is only by removing the effects of the remainder in some way that the connection can be established. Suppose x_1, x_2, x_3 are measured (see figure 6) and it is suspected that x_1 and x_2 cause x_3, i.e., (linear) systems S_1, S_2 are believed to link x_1, x_2 to x_3. Suppose further that in reality link S_2 is not present (but this is unknown) and also that both x_1 and x_2 are themselves 'caused' by some fundamental (unmeasured) process θ ensuring that x_1 and x_2 are highly coherent. If the ordinary coherence function $\gamma^2_{x_2y}$ is formed this will be large tending to 'confirm' the belief in the presence of path S_2. But if the effect that x_1 may have on both x_2 and y is removed before trying to establish a link between x_2 and y then it would soon be realised that no such link existed. The procedure for effecting this 'removal' is now explained.

Consider a two input single output situation as in fig. 7 (n may be measurement noise or introduced to model nonlinearities) and the objective might be to (i) establish the relative magnitudes of noise to 'linear effects' at the output, (ii) to find the relative importance of inputs x_1, x_2, (iii) to estimate the transfer functions H_1, H_2. (It is assumed that the ordinary coherence function $\gamma^2_{x_1x_2}$ is not unity since in this case it is only a single input problem.)

As a starting point it is convenient to decompose the processes involved as shown in fig. 8. x_2 and x_3 are split into two components each where y_2 and y_4 are fully coherent with y_1 (i.e., those parts of x_2 and x_3 accounted for by optimal linear operations on x_1 through filters L_1, L_2). y_3 is uncoherent with y_1, and is a residual random variable denoted $x_{2.1}$ (similarly for $x_{3.1}$). From these residual random variables one can define residual spectra which are ordinary spectra formed from residual variables and partial coherences which are ordinary coherences formed from residual variables. Referring in particular to 'stage 1' in the figure it is noted that x_2 and x_3 have been conditioned with respect to x_1 (this choice is arbitrary) and from the earlier work one can write down the following:- (i) The 'coherent output power', i.e., that proportion of the power of x_2 accounted for by linear operations on x_1 is $\gamma^2_{12}S_{22}$. (N.B. $S_{22} = S_{x_2x_2}$, etc.) (ii) The 'noise' power $S_{y_3y_3}$ is a residual power spectral density written $S_{22.1}$ and is $S_{22.1} = (1 - \gamma^2_{12})S_{22}$. (iii) The optimal linear filter L_1 is S_{12}/S_{11}. Identical expressions hold for processes x_1, x_3 with 3 replacing 2 above. It is important to realise that L_1 and L_2 are not H_1 and H_2. In fact since $x_{2.1}$ and $x_{3.1}$ denote what is left after the linear effect of x_1 is removed from x_2 and y it should be clear that $H_2 = S_{23.1}/S_{22.1}$ (to find H_1 reverse the roles of 1 and 2). It only remains to obtain an expression for the residual cross spectrum $S_{23.1}$ in terms of x_1, x_2, x_3. Since

$$S_{23.1} = \lim_{T\to\infty} E\left[\frac{1}{T}(X^*_{2.1}(f)X_{3.1}(f))\right]$$

then using the signal flow algebra of the figure and expressions for L_1, L_2 gives $S_{23.1} = S_{23} - (S_{21}S_{13}/S_{11})$ i.e., the residual spectral density is evaluated in terms of ordinary spectra.

These concepts introduced at 'stage 1' can be extended in similar fashion to 'stage 2' where x_3 can be further conditioned, now with respect to $x_{2.1}$. y_6 is that part of $x_{3.1}$ coherent with $x_{2.1}$ and y_7 is that part of x_3 that is uncoherent with both x_1 and x_2 or in other words that part of x_3 not accounted for by linear operations on x_1 and x_2. This means y_7 is the residual random variable formed by conditioning x_3 with respect to both x_1 and x_2 and is written $x_{3.12}$. At stage 2 (by analogy with stage 1) one can write down the following:- (i) The output noise power spectral density is $S_{33.12} = (1 - \gamma^2_{23.1})S_{33.1}$ where (ii) The partial coherence function $\gamma^2_{23.1}$ is $\gamma^2_{23.1} = S_{23.1}S_{32.1}/S_{22.1}S_{33.1}$ and (iii) The optimal filter L_3 is $S_{23.1}/S_{22.1}$. The total output power spectral density can be decomposed as the sum of three terms (and this can be done directly from the diagram)

$$S_{23} = \gamma^2_{13}S_{33} + \gamma^2_{23.1}S_{33.1} + S_{33.12}$$

part fully coherent with x_1 | part fully coherent with x_2 after x_1 has been removed from x_2 and x_3 | uncoherent with both x_1 and x_2

The multiple coherence function (written $\gamma^2_{y.x}$) is defined by analogy with the ordinary coherence function as for a single input system, i.e., it is that fraction of power accounted for in the output via linear relationships between input and output, i.e., $\gamma^2_{y.x} = (S_{33} - S_{33.12})/S_{33}$ and by the expressions above can be written $\gamma^2_{y.x} = 1 - (1 - \gamma^2_{13})(1 - \gamma^2_{23.1})$.

It is appropriate to emphasise that the multiple coherence function is a measure of how well the two inputs x_1 and x_2 account for the measured response of the system, whilst the partial coherence is a measure of how well an additional signal (in this case x_2) improves the predicted output.

All the formulae above can be related back to the direct and cross spectra computed from the original three signals and can be applied directly to higher order systems to recover the general formulae given by Bendat [17] by simply adding on more 'stages' to the figure for appropriately more inputs and designating higher order residual processes in the obvious way (e.g., $x_{4.123}$). With reference to the computation of the transfer functions H_1, H_2, an expression for H_2 has been written down earlier, but the H_i are easily expressed in terms of the L_j (themselves easily written down). By simply following the signal flow algebra in the figure one can write $H_2 = L_3$ and $H_1 = L_2 - L_3L_1$. It is appropriate to note (see the discussion following reference [15], that the manipulations to evaluate the transfer functions amount to a Gauss elimination solution of the input-output cross spectral equations but that the approach as outlined here at least offers some physical interpretation in the unravelling of complex problems. Finally it is pointed out that Bendat [17] has given a comprehensive summary of the statistical properties of residual processes.

4. FUNCTIONS OF MORE THAN ONE VARIABLE

In the foregoing the data to be analysed have been assumed to be real valued functions of a single variable. Often data may be regarded as functions of more than one variable (e.g., space and time) and the methods already outlined may be generalised to include such phenomena. Two (and more) dimensional signal processing occurs in many fields with many applications, e.g., image processing, seismic signal analysis, beamforming, etc. and a few examples relevant to vibration and acoustic problems are noted later. To begin with, however, a brief general description of two dimensional signal processing in both continuous and discrete form is given.

4.1 Two Dimensional Signal Analysis

Continuous space/time processes

Let $s(x, t)$ denote a signal that is a function of both time t and x a space variable. Obviously this could be generalised to include more space variables or 'simplified' to exclude time. Discussion of the form chosen is sufficient for our purposes. Assume $s(x, t)$ to be a deterministic function of both x and t; we can define a two dimensional Fourier transform

$$S(k, \omega) = \int_{-\infty}^{\infty}\int s(x, t)e^{-j(kx+\omega t)}dxdt \qquad (24)$$

By regarding $S(k, \omega)$ as the result of two one-dimensional transforms [18] it is simple to establish the inversion formula

$$s(x, t) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int S(k, \omega)e^{j(kx + \omega t)}dkd\omega \qquad (25)$$

The two dimensional Fourier transform $S(\omega, t)$ is a function of the angular frequency ω (radians/second) and the wavenumber k (radians/metre) describing the harmonic structure of $s(x, t)$ in terms of distance. The (temporal) frequency component ω has period $T = 2\pi/\omega$ seconds and the (spatial) wavenumber component k has wavelength $\lambda = 2\pi/k$ metres. (We note that both radians/second and Hz are used in this section as appropriate.)

Input-output relationships for functions of two variables may be developed. In fig. 9 the impulse response function of a generalised linear system operating both in time and space is $h(x, t)$ and the output variable is $y(x, t)$, so that

$$y(x, t) = \int_{-\infty}^{\infty}\int h(x - x_1, t - t_1)s(x_1, t_1)dx_1dt_1.$$

It can be verified [18] that $Y(k, \omega) = H(k, \omega)S(k, \omega)$. *Stochastic processes* may be included in the above formulation. Suppose $s(x, t)$ is a random process defined such that the process is *stationary* in time and *homogeneous* in space, then $E[s(x_1,t_1)s(x_2,t_2)] = R(x_2 - x_1;\ t_2 - t_1)$ or $E[s(x,t)s(x + \xi, t + \tau)] = R_{s_1s_2}(\xi;\tau)$. The cross spectral density $S_{s_1s_2}(k, \omega)$ is defined

$$S_{s_1s_2}(k, \omega) = \int_{-\infty}^{\infty}\int R_{s_1s_2}(\xi; \tau)e^{-j(k\xi+\omega\tau)}d\xi d\tau$$

which has a corresponding inversion formula. Papoulis [18] discusses these aspects further.

Discrete space/time processes

Since processing is generally done using digital methods it is appropriate to reconsider (briefly) the above Fourier relationships using discretized signals. A more detailed account may be found in the chapter by Peardon [19] or Rabinerand Gold [20]. A two dimensional signal may be inherently discrete (at least in one variable) or generated by sampling a function of continuous variables. It is convenient to regard the result as a two dimensional sequence (array), say, $s(m\Delta x, n\Delta t)$, . The sampling is assumed uniform in both the space and time domains. Ignoring the sample separations Δx, Δt we can define the *two dimensional Z transform* as

$$S(z_1, z_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} s(m, n) z_1^{-m} z_2^{-n}$$

If this is evaluated on the unit circle, i.e., $z_1 = e^{j\omega_1}$; $z_2 = e^{j\omega_2}$ we obtain the Fourier transform relationship

$$S(e^{j\omega_1}, e^{j\omega_2}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} s(m, n) e^{-j\omega_1 n} e^{-j\omega_2 n}$$

with the inverse

$$s(m, n) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} S(e^{j\omega_1}, e^{j\omega_2}) e^{j\omega_1 n} e^{j\omega_2 n} \, d\omega_1 d\omega_2$$

(The notation adopted here is that appropriate to discrete sequences.) The *two dimensional discrete Fourier transform (DFT)* is the natural progression from the above so that the transforms can be computed using digital methods. So, if $s(m, n) = 0$ for $0 \leq m \leq M - 1$, $0 \leq n \leq N - 1$, then

$$S(k, \ell) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} s(m, n) e^{-j2\pi(\frac{km}{M} + \frac{\ell n}{N})}$$

The inverse is

$$s(m, n) = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{\ell=0}^{N-1} S(k, \ell) e^{j2\pi(\frac{km}{M} + \frac{\ell n}{N})}$$

Fast Fourier transform methods [19] may be employed to reduce the computational effort to evaluate these transforms.

Two dimensional digital filters can be constructed and the effect of operation on array $s(m, n)$ written as

$$y(m, n) = \sum_{k} \sum_{\ell} h(m - k, n - \ell) s(k, \ell)$$

where $h(m, n)$ is the response of the filter to an impulse at $m = n = 0$. The filter frequency response is the double Fourier transform of $h(m, n)$. Two dimensional digital filters may be designed and implemented using methods that are generalisations of the appropriate one dimensional methods mentioned in the first lecture. It is important to note however that the DFT $S(k, \ell)$ of an array of numbers $s(m, n)$ is *periodic* in both k and ℓ and that this periodicity must be borne in mind if it is intended to implement *linear* convolution using frequency domain methods, i.e., the 'distortion' that can result from the effects of cyclic or periodic 2-D convolution using FFT methods must be accounted for [19].

Aliasing for two-dimensional signals is an aspect that requires special consideration as it can have a very marked effect on the appearance of Fourier transformations. Recall that for a one dimensional process a sampled signal may be regarded as the product of the continuous wave and a delta 'comb', i.e., $s(n\Delta t) = s(t) \sum_{n} \delta(t - n\Delta t)$ so that $S(e^{j2\pi f}) = S(f) * \sum_{k} \delta(f - k/\Delta t)$ (* denotes convolution). This argument may be extended to two dimensions. First generalise the delta 'comb' to a delta 'brush' and write

$$s(m\Delta x, n\Delta t) = s(x, t) \sum_{m} \sum_{n} \delta(x - m\Delta x, t - n\Delta t)$$

This results in a Fourier transform

$$S(e^{j2\pi f_1}, e^{j2\pi f_2}) = S(f_1, f_2) * \sum_{k} \sum_{\ell} \delta(f_1 - \frac{k}{\Delta x}, f_2 - \frac{\ell}{\Delta t})$$

This equation indicates that the continuous transform $S(f_1, f_2)$ is 'copied' onto rectangular cells spaced at intervals $1/\Delta x$ and $1/\Delta t$ in the f_1, f_2 directions. This results in a two dimensional version of the overlapping that is the result of aliasing, thereby complicating the appearance of the transform evaluated within a single cell (see ref. [19]).

4.2 Applications

Space-time functions occur in various contexts and we shall very briefly indicate a few here.

Rough ground profiles

The irregularities of a rough surface may be written as a function of two space variables, i.e., $s(x, y)$ where now time plays no role. Such height profiles are useful when describing the response of vehicles running over rough terrain. If the surface is assumed random and homogeneous then wavenumber spectra can be computed from the appropriate correlation function. Numerous models for rough ground wavenumber spectra have been proposed in the literature (see, for example, [21]).

Random pressure fluctuations

Often interest is in the characteristics of a pressure field in some spatial domain where the instantaneous value is $p(\underline{r}, t)$ where $\underline{r}$ is the vector position of the point in space. Spatial cross correlation and cross spectra arise in the description of the structure of pressure fields in turbulent flow (e.g., [22]).

Seismic signal processing [19]

Seismic 'sections' are obtained by placing in juxtaposition sets of one dimensional records that are obtained by probing some feature of interest with an acoustic excitation. Each record may be obtained by triggering a response at differing spatial positions (see fig. 10).

It is only by placing such records together and viewed as a whole that the geological structure becomes apparent. It is pointed out that the time record may be continuous but spatial sampling is built in automatically.

Beam forming [23,24]

The problem of estimating the bearing of some source using an array of sensors is often referred to as 'beam forming'. Many array signal processing systems use delay and sum beam forming to estimate source bearings assuming that the process is nondispersive at the array. Furthermore, beam forming can be directly related to frequency-wavenumber Fourier analyses. This is now explained following Hinich [24]. As a simple model consider (fig. 11) a linear array of M sensors receiving coherent radiation (in this case a single frequency plane wave).

If a_o is the wave velocity and θ_o the angle between the direction of propagation and the array axis then the signal at the ℓ^{th} sensor may be expressed (if noise is zero) as

$$s(t, x_\ell) = Ae^{j\omega_o(t - \frac{x_\ell \cos\theta}{a_o})} \qquad \text{(A complex)}$$

Since each of the M signals is merely a delayed copy of the preceding one, if the appropriate delay was built in to each and the signals added the signals would reinforce one another to produce a signal with 'maximum power'. If we assume that the direction of arrival of the wave is θ and form

$$B(t, \theta) = \sum_{\ell=1}^{M} s(t + \tau_\ell, x_\ell)$$

where $\tau_\ell = x_\ell \cos\theta/a_o$, then $B(t, \theta)$ has maximum power when $\theta = \theta_o$. Note that $B(t, \theta)$ can be written

$$B(t, \theta) = \sum_{\ell=1}^{M} s(t, x_\ell)e^{jkx_\ell} \qquad (k = \frac{\omega_o}{a_o}\cos\theta)$$

and so beamforming is equivalent to computing the spatial Fourier transform of the M signals from the array. Aliasing effects are discussed by Hinich in [23]. These ideas may be generalised to include noise, broadband signals and larger arrays. There is a lot of interest now in adaptive beam forming [25] and the Maximum Entropy method has also been applied [26].

Source location

A collection of methods commonly referred to as 'source location techniques' have been described in a review paper by Fisher and Glegg [27]. Additional detailed material may be found in [28]. The objective of these methods is to establish the relative importance (in the far field) of a collection of sound sources that occur simultaneously. Whilst the techniques developed may, in principle, be applied to a variety of situations, a particular application that has received considerable attention is the evaluation of relative source strengths for noise radiated from a jet engine. It is noted that in contrast to 'usual' radar and sonar applications there is considerable *a-priori* knowledge of the system under study. More specifically, noise is radiated from the engine inlet, the bypass duct exhaust, the hot core exhaust and also the jet flow. Any engine noise reduction study would require information concerning the relative importance of each component as far as the far field noise level is concerned. Several techniques have been proposed to deal with the problem and a brief comparative assessment is given in [27]. All the methods involve the problems of space-time signal processing and some show marked similarities with 'beam forming' discussed earlier. For brevity only the polar correlation [28] technique will be outlined below.

Some assumptions/features of the technique are listed below.

(a) The method uses an array of microphones on a polar arc centred at a point in the source region (see fig. 12). Equivalently a reference microphone and a traversing second microphone may be used.

(b) In what follows below we follow [27] and assume a line source of (correlated) monopoles.

The signal at a microphone at angular position α in the far field is

$$p_2(t) \simeq \int_y \frac{1}{(r_o - y\sin\alpha)} F(y, t - \frac{(r_o - y\sin\alpha)}{a_o})dy$$

where F is the source strength at y of an equivalent distribution of monopoles on the jet axis which if r_o is sufficiently large is approximately

$$p_2(t) \simeq \frac{1}{r_o}\int_y F(y, t - \frac{(r_o - y\sin\alpha)}{a_o})dy$$

The reference microphone signal ($\alpha = 0$) is

$$p_1(t) = \frac{1}{r_o} \int_{y'} F(y', t - \frac{r_o}{a_o}) dy'$$

If the cross correlation function between microphones is $R_{12}(\tau) = E[p_1(t)p_2(t + \tau)]$ then

$$R_{12}(\tau) = \frac{1}{r_o^2} \int_y \int_{y'} R_{FF}(y', y, \frac{\tau + y \sin \alpha}{a_o}) dy'dy$$

where

$$R_{FF}(y', y, \frac{\tau + y \sin\alpha}{a_o}) = E\{F(y', t - \frac{r_o}{a_o})F(y, t + \tau - \frac{r_o}{a_o} + \frac{y \sin\alpha}{a_o})\}.$$

The cross spectral density $S_{12}(\omega)$ is

$$S_{12}(\omega) = \int_y e^{j\omega \frac{y \sin \alpha}{a_o}} S(y, \omega)\, dy$$

where

$$S(y, \omega) = \frac{1}{2\pi r_o^2} \int_{-\infty}^{\infty} R_{FF}(y', y, \theta) e^{-j\omega\theta}\, d\theta dy'.$$

The definition of source strength is chosen as $S(y, \omega)$ and so this may be found from $S_{12}(\omega)$ as

$$S(y, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{12}(\omega) e^{-j\omega \frac{y \sin \alpha}{a_o}} d(\frac{\omega \sin \alpha}{a_o})$$

Thus, in principle at least, the source strength may be computed from knowledge of $S_{12}(\omega)$.

However some difficulties should be immediately apparent, namely, (a) the limits for the last integral are $\pm\omega \sin \alpha_{max}/a_o$ (and in any event no greater than $\pm \omega/a_o$). This imposes a resolution limit. (b) The cross spectral density is not known at a continuum of angular positions but practical considerations dictate a finite set, then this spatial discretization results in aliased images of the sources.

Detailed discussions of these problems including the use of a graded microphone array (i.e., non-uniform spacing) is given in full in the thesis by Glegg [28].

REFERENCES

1. K.G. Beauchamp. Signal Processing. George Allen and Unwin Ltd. 1973.
2. J.S. Bendat and A.G. Piersol. Random Data: Analysis and Measurement Procedures. Wiley Interscience. 1971.
3. R.K. Otnes and L. Enochson. Applied Time Series Analysis Vol. 1. Basic Techniques. Wiley and Sons. 1978.
4. G.M. Jenkins and D.G. Watts. Spectral Analysis and its Applications. Holden Day. 1968.
5. C. Bingham, M. Godrey and J.W. Tukey. Modern techniques of power spectrum estimation. IEEE Trans. Audio and Electro-acoustics, vol. AU-15, No. 2. 1967.
6. G.C. Carter. Bias in magnitude-squared coherence estimation due to misalignment. IEEE Trans. on Acoustics, Speech and Signal Processing. Vol. ASSP-28, No. 1. 1980.
7. C.M. Rader. An improved algorithm for high speed autocorrelation with application to spectral estimation. IEEE Trans. Audio and Electroacoustics, Vol. AU-18. 1970.
8. P.D. Welch. The use of the fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans. Audio and Electroacoustics, Vol. AU-15. 1967.
9. P.E. Wellstead. Nonparametric methods of system identification in 5th IFAC Symposium on Identification and System Parameter Estimation. Ed. R. Isermann, held at Darmstadt. Pergamon Press. 1979.
10. C.K. Yuen. Comments on modern methods for spectrum estimators. IEEE Trans. Acoustics, Speech and Signal Processing. Vol. ASSP-27, no. 3. 1979.
11. A.P. Sage. Optimum Systems Control. Prentice Hall Inc. 1968.
12. J.S. Bendat. System identification from multiple input/output data. Journal of Sound and Vibration 49(3). 1976.
13. J.S. Bendat. Solutions for the multiple input/output problem. Journal of Sound and Vibration 44(3). 1976.
14. J.S. Bendat. Statistical errors in measurement of coherence functions and input/output quantities. Journal of Sound and Vibration 59(3). 1978.
15. J.S. Bendat. Stochastic Problems in Dynamics. Ed. B.L. Clarkson. Procedures for frequency decomposition of multiple input/output relationships. Pitman Pub. Ltd. 1977.
16. C.J. Dodds and J.D. Robson. Partial coherence in multivariate random processes. Journal of Sound and Vibration 42. 1975.

17. J.S. Bendat and A.G. Piersol. Engineering Applications of Correlation and Spectral Analysis. John Wiley and Sons. 1980.

18. A. Papoulis. Systems and Transforms with Applications in Optics. McGraw Hill. 1968.

19. L.G. Peardon. Two dimensional discrete signal processing in Applications of Time Series Analysis. Course Notes. ISVR. Southampton University.

20. L.R. Rabiner and B. Gold. Theory and Application of Digital Signal Processing. Prentice Hall. 1975.

21. J.D. Robson. Road surface description and vehicle response. J. of Vehicle Design. Vol. 1 No. 1. 1979.

22. P.O.A.L. Davies and C.A. Mercer. Phase velocity measurements using the cross spectrum. Int. Symposium on Measurement and Process Identification. Bradford, England. Jan. 1973.

23. M.J. Hinich. Processing spatially aliased arrays. Jnl. Acoustical Society of America, 64(3). 1978.

24. M.J. Hinich. Frequency-wavenumber array processing. Journal of the Acoustical Society of America, 69(3). 1981.

25. Wen-Wu Shen. A constrained minimum power adaptive beam former with time varying adaptation rate. Geophysics, Vol. 44, No. 6. June 1979.

26. R.N. McDonough. Maximum entropy spatial processing of array data. Geophysics, Vol. 39, pp. 843-851. 1974.

27. M.J. Fisher and S.A.L. Glegg. Advanced Course on Noise and Vibration: Course notes. ISVR, Southampton University. Source location.

28. S.A.L. Glegg. Ph.D. Thesis, University of Southampton. Jet Noise Source Location. 1979.

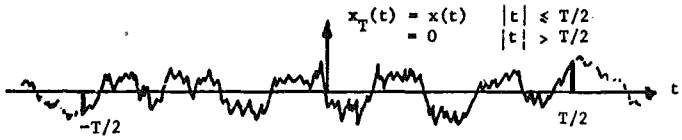


Fig. 1. Sample function of a random process.

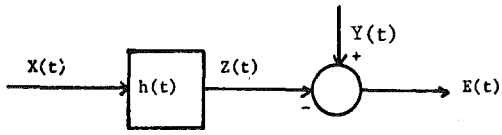


Fig. 2. Optimal estimation of h(t) (to relate X and Y)

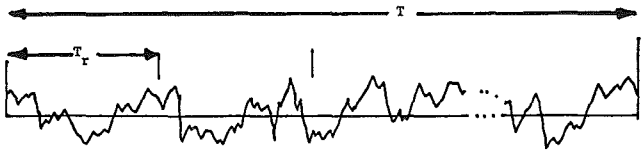


Fig. 3. Segmenting of a time history.

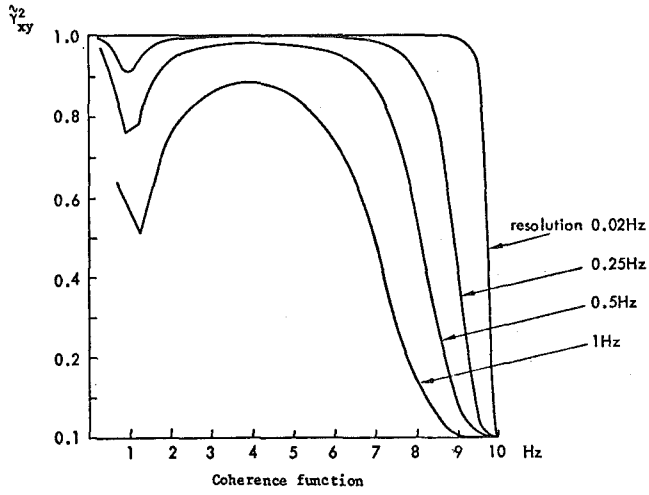


Fig. 4. The effect of changing resolution on the coherence estimate.

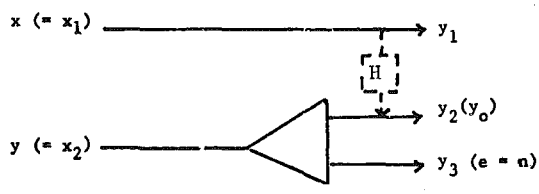


Fig. 5. A single input (x) single output (y) system: y_3 is a residual random variable.

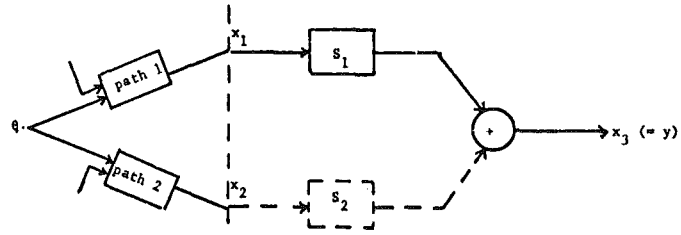


Fig. 6. An apparent two input (x_1, x_2) single output (x_3) system.

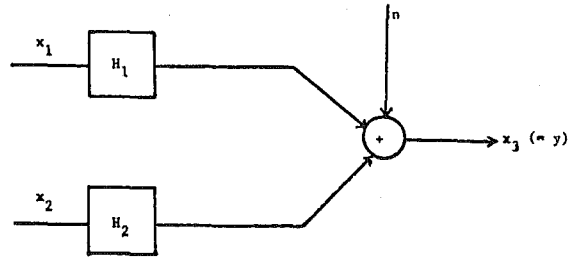


Fig. 7. A two input single output system including noise.

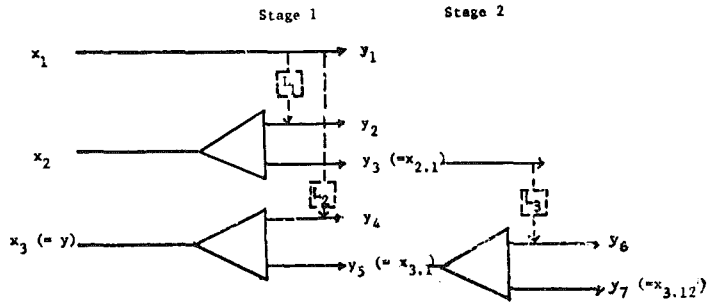


Fig. 8. Decomposition of a two input single output system in terms of residual spectra.

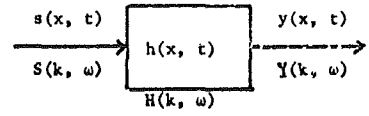


Fig. 9. A two dimensional filter.

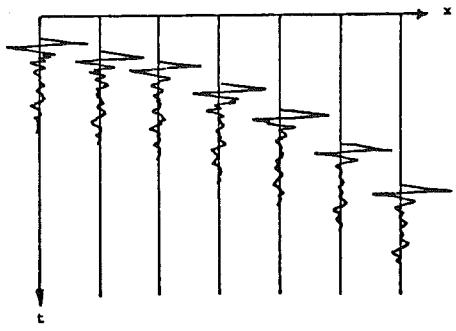


Fig. 10. A seismic section.

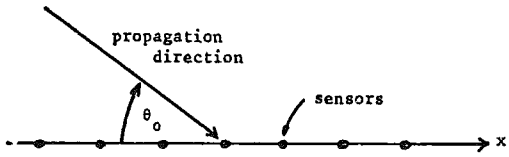


Fig. 11. A linear array of sensors receiving coherent radiation.

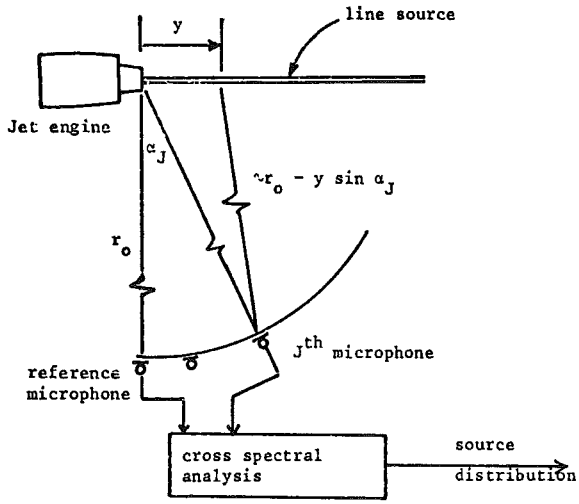


Fig. 12. The polar correlation method of source location.

NON-STATIONARITY AND NONLINEARITY IN DATA ANALYSIS

J.K. Hammond

Institute of Sound and Vibration Research
University of Southampton
Southampton SO9 5NH
England

SUMMARY

Problems involving nonstationary random processes and nonlinear systems pose urgent and challenging questions. In what is a wide area of study we shall, in this chapter, merely select a few aspects in an attempt to highlight some of the approaches to the problems that arise. Both time and frequency domain methods are considered, with emphasis on procedures that are generalisations of methods that are applicable to stationary processes and linear systems. Whilst the title explicitly mentions 'data analysis' the treatment is broader since the problem of modelling of processes is also considered as this may be the basis of any subsequent analysis procedure.

The work is broken in three portions. The first deals with nonstationary random processes and begins with some general considerations related to analysis of nonstationary data and then outlines a time domain approach to the modelling of vehicles travelling over rough ground at variable speed. This is followed by a quite different description of a class of nonstationary phenomena using the evolutionary spectral density. The second part deals with nonlinear phenomena and begins again in the time domain with a brief survey of a body of theory devoted to the solution of randomly driven nonlinear systems. and concludes with a description of a frequency domain approach for nonlinear systems using the bispectrum. The final section of the chapter explains a signal processing approach that might be called 'nonlinear' (rather than the phenomenon from which the data arises). This is homomorphic signal processing which has very wide application, one of which is the treatment of echo-like processes.

1. NONSTATIONARY RANDOM PROCESSES

A stationary process is one for which the probabilistic structure is invariant under a shift in time, though for 'weak' stationarity this requirement is reduced to include only the first and second order probability density functions. A nonstationary process is one that contravenes the former condition, though we may decide only to refer to the latter (weaker) condition. In view of this we may say that the nonstationarity of a random process $x(t)$ is embodied (up to second order properties) in the time varying mean $(\mu_x(t))$, the time varying variance $(\sigma_x^2(t))$ and the autocorrelation function $R_{xx}(t_1, t_2)$, which is no longer merely a function of lag $(t_2 - t_1)$. A further consequence of nonstationarity is that the notion of the power spectral density is no longer admissible (at least not without careful redefinition - see section 1.2).

As far as the analysis of nonstationary data is concerned it seems evident that numerous realisations of the same phenomenon are required and then estimates of the required quantities may be computed by ensemble averaging. For example, an estimate of the mean at time t from M records would be $\bar{x}(t) = \frac{1}{M}\sum_{i=1}^{M} x_i(t)$. It is often the case, however, that suitably large numbers of realisations are unavailable, indeed we may have only a single record of the process and are confronted with the problem of analysing the data. The only procedure open to us is to create estimates of the required quantities at time t, say, from data 'in the vicinity of' that time, for example. We may visualise using a 'window' centered on the time t, having a limited extent in both the t^+ and t^- directions (perhaps with taper) defining the 'local' data set to be used in the computation. Such local averages may easily be computed, but the difficulty is deciding on their usefulness. To obtain estimates that do not suffer large sample fluctuations means in turn that the underlying nonstationary trend is often smeared out. Clearly as much prior information concerning the structure of the nonstationarity should be incorporated and accounted for if this is to be an effective procedure. It may happen though that prior information is lacking and that it might be that the issue is as basic as whether the process is stationary or not. This raises the question of the development of tests for (non) stationarity and some have been proposed in the literature. Bendat and Piersol [1] consider criteria aimed at deciding whether a sequence of, say, mean square values (computed from independent data segments) display some underlying trend. More recently, Subba Rao and Tong [2] have considered tests for nonstationarity based upon evolutionary spectra.

However, if prior information allows a model of the process to be constructed then a simple pragmatic approach to analysis can be effective. Suppose, for example, it is known that a nonstationary process $z(t)$ is a uniformly modulated process, i.e., $z(t) = m(t)x(t)$ where $x(t)$ is a zero mean stationary process and $m(t)$ a nonrandom 'slowly' varying function. Since $\sigma_z^2(t) = m^2(t)\sigma_x^2$ a 'running' estimate of $\sigma_z^2(t)$ leads directly to an estimate of $m(t)$, say $\tilde{m}(t)$ (any convenient envelope detection procedure might do). Next the portion $x(t)$ can be estimated and analysed as a stationary process allowing the nonstationary properties of $z(t)$ to be written, e.g.,

$$\hat{R}_{zz}(t_1, t_2) = \tilde{m}(t_1)\tilde{m}(t_2)\hat{R}_{xx}(t_2 - t_1)$$

(we note that $m(t)$ can only be estimated to within a scale factor).

It is more usually the case that the nonstationarity is not 'separable' as above but more complicated. In the next section we outline a vibration problem where this arises and propose a time domain model to describe it. The objective is to demonstrate a model for the prediction of nonstationarity and will not consider the analysis of data arising from realisations of the process.

1.1 The Nonstationary Response of Vehicles Travelling Over Rough Ground [3]

The analysis of the behaviour of vehicles traversing uneven ground profiles has been considered by many workers over a number of years and a common assumption is that the rough ground is a sample from a homogeneous (spatially stationary) random process, and following from this it is straightforward to predict the stationary response of a vehicle modelled by linear dynamics travelling over the ground at constant velocity. If the velocity of the vehicle is variable this results in nonstationary response and in ref. [3] a novel state space approach for the solution of the problem for linear dynamics is given, and in [4] nonlinear dynamics are included using the method of equivalent linearisation (see section 2). We shall indicate the approach used in [3].

Figure 1 is a simple model of a vehicle running over rough ground. (The equations may be developed for more complicated dynamics but the method is best explained using the simple model.) The vehicle travels at (variable) speed $\dot{s}$ over the ground whose height $h(s)$ is characterised by the spatial autocorrelation function $R_{hh}(s_2 - s_1) = E[h(s_1)h(s_2)]$. If s is written as a function of time (i.e., $s(t)$) to denote vehicle motion an equivalent temporal autocorrelation function is $R_{hh}(t_1, t_2) = R_{hh}(s(t_2) - s(t_1))$ which now (for general $s(t)$) describes a nonstationary process. If $g(\tau)$ is the impulse response function of the vehicle and $y(t)$ the (say displacement) response then the response autocorrelation function is obtained from

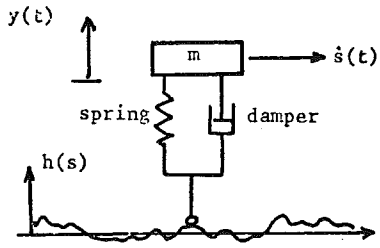


Fig. 1

$$R_{yy}(t_1, t_2) = \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} g(t_1 - \tau_1)g(t_2 - \tau_2)R_{hh}(\tau_1, \tau_2)\,d\tau_1 d\tau_2$$

Reference [3] contains a brief summary of solutions based on this equation and others that have appeared in the literature.

A key assumption in the use of the state space approach described in detail in [3] is that the ground surface undulation may be described as the output of a white noise excited shaping filter (in the spatial domain). This allows the linking of the differential equations for the dynamics and the excitation into a composite state form. The vehicle dynamics for the simple situation depicted in fig. 1 are described (assuming linear dynamics) in the usual way as

$$\ddot{y} + 2\zeta\omega_o\dot{y} + \omega_o^2 y = 2\zeta\omega_o\dot{\tilde{h}} + \omega_o^2\tilde{h} \qquad (1)$$

where $\tilde{h}$ is the notation for the height profile regarded as a function of time, i.e., $\tilde{h}(t) \equiv h(s(t))$. We shall now assume that the spatial autocorrelation function of the height is

$$R_{hh}(\xi) = \sigma^2 e^{-\alpha|\xi|} \qquad (\xi = s_1 - s_2). \qquad (2)$$

The height h represented in this form may be modelled as the output of a first order shaping filter which in differential equation form is (using space variable s)

$$\frac{dh}{ds} + \alpha h = Kw(s) \qquad (3)$$

K is the strength of the white (in space) process selected to be $\sigma\sqrt{2\alpha}$ where σ is the standard deviation of h. (This choice of K ensures the correct variance of h.) The idea now is to link (3) with (1) but we first need to use a common independent variable, either s or t. We shall use t noting that $d(\,)/ds \equiv (1/\dot{s})/(d(\,)/dt)$ so that (3) may be rewritten as

$$\frac{d\tilde{h}}{dt} + \dot{s}\alpha\tilde{h} = \dot{s}Kw[s(t)]. \qquad (4)$$

Now we couple (4) with (1) into the state form

$$\frac{d}{dt}\begin{bmatrix} y \\ \dot{y} \\ \tilde{h} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ \omega_o^2 & -2\zeta\omega_o & (\omega_o^2 - 2\zeta\omega_o\alpha\dot{s}) \\ 0 & 0 & -\alpha\dot{s} \end{bmatrix}\begin{bmatrix} y \\ \dot{y} \\ \tilde{h} \end{bmatrix} + \dot{s}\begin{bmatrix} 0 \\ 2\zeta\omega_o K \\ K \end{bmatrix} w(s(t)) \qquad (5)$$

As is shown in [3], if a state system is described by $\dot{\underline{x}} = \underline{A}\underline{x} + \dot{s}\underline{B}\underline{w}[s(t)]$ with $E[\underline{w}(s_1)\underline{w}^T(s_2)] = Q\delta(s_1 - s_2)$ then the zero lag covariance matrix $P(t) = E[\underline{x}(t)\underline{x}^T(t)]$ satisfies $\dot{P} = AP + PA^T + \dot{s}BQB^T$ and so knowledge of A, B, $\dot{s}$ and Q ensures that the response statistics can be computed even in a nonstationary situation by simply integrating the equation for P numerically. In the example above it means that the response of the vehicle may be obtained for any velocity profile and solutions are presented in [3] for constant acceleration and sinusoidally varying velocities. As mentioned earlier this technique can be extended to include nonlinear dynamics if the method of (stochastic) equivalent linearisation is used for the nonlinear elements (see section 2). It is emphasised though that the excitation must admit a (spatial) shaping filter representation in order that this technique be applicable.

1.2 Evolutionary Spectra

In contrast to the time domain approach to treating nonstationary processes it is interesting to consider the possibility of applying frequency domain methods since spectra have a physical appeal and the notion of a 'time varying spectral density' is intuitively attractive. Many forms of two dimensional

spectra have been defined for nonstationary processes but one which retains the physical interpretation of the ordinary spectral density is the 'evolutionary spectral density' introduced by Priestley [5]. This was later generalised to include cross spectra [6] and utilised for a vibration application in [7]. The object of this section is to explain the basis of the method.

A stationary random process has a generalised Fourier representation

$$x(t) = \int_{-\infty}^{\infty} e^{j\omega t}\, dZ(\omega) \qquad (6)$$

(This Stieltjes integral reduces to the usual form when $x(t)$ is a transient so that $\frac{dZ(\omega)}{d\omega} = X(\omega)$, say.) Since $x(t)$ is random $Z(\omega)$ changes from realisation to realisation, i.e., $dZ(\omega)$ is itself a random variable and for a *stationary process* the increments $dZ(\omega_1)$, $dZ(\omega_2)$ are uncorrelated for $\omega_1 \neq \omega_2$, i.e., $E[dZ^*(\omega_1)dZ(\omega_2)] = 0$, $\omega_1 \neq \omega_2$.

This property ensures that the autocorrelation function $E[x(t)x(t+\tau)]$ is a function of lag τ only and can be written

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} e^{j\omega\tau}\, E\{[dZ(\omega)]^2\}$$

which can in turn be written $R_{xx}(\tau) = \int_{-\infty}^{\infty} e^{j\omega\tau} S_{xx}(\omega)\, d\omega$ and this introduces the spectral density $S_{xx}(\omega)$. The key features in this development are the representation of $x(t)$ as the sum of sines and cosines and the orthogonality of the increments $dZ(\omega)$. If $x(t)$ is *nonstationary* some processes still admit representation (6) but the increments $dZ(\omega)$ are no longer orthogonal and this results in two dimensional (bi-frequency) spectra written $S(\omega_1, \omega_2)$ which are difficult to interpret. Priestley's objective was to define a two dimensional spectral density where the dimensions are frequency and time written $S_{xx}(t, \omega)$, say, such that for a fixed t, $S_{xx}(t, \omega)$ does indeed describe the local power distribution over frequency, so that $S_{xx}(t, \omega)$ has exactly the same *interpretation* as the power spectral density of a stationary process. The starting point for this is equation (6). It is necessary to develop a representation for $x(t)$ still retaining the notion of 'frequency' and the orthogonality property and yet relating to nonstationarity. Priestley chose the basic elements in the representation as modulated sines and cosines $A(t, \omega)e^{j\omega t}$ so that

$$x(t) = \int_{-\infty}^{\infty} A(t, \omega) e^{j\omega t} dZ(\omega) \qquad (7)$$

where $Z(\omega)$ is still an orthogonal process. Not all nonstationary processes may be represented in this way but if there exist a family of 'oscillatory functions' $\{A(t, \omega)e^{j\omega t}\}$ in terms of which $x(t)$ can be represented as in (7) then $x(t)$ is said to be an 'oscillatory process'. A useful interpretation of equation (7) is that such a representation arises naturally if the nonstationary process may be regarded as the output of a time varying filter driven by a stationary process. In that case the $A(t, \omega)$ function can take a variety of forms but in the original papers conditions were imposed on $A(t, \omega)$ to ensure that it was 'slowly varying'. It is a simple matter to use (7) and the orthogonality condition to show that

$$E[x^2(t)] = \int_{-\infty}^{\infty} |A(t, \omega)|^2 S_{xx}(\omega)\, d\omega \qquad (8)$$

so that the evolutionary spectral density is $S_{xx}(t, \omega) = |A(t, \omega)|^2 S_{xx}(\omega)$ and from (8) is seen to satisfy the required interpretation as a power decomposition.

Estimation of evolutionary spectra

We now consider the problem of estimating the evolutionary spectral density from a single realisation. The procedure proposed is roughly equivalent to splitting the process up into segments and estimating the spectral content within each interval and there are variants of the basic method. We shall outline the procedure described by Priestley using complex demodulation and which is summarised in the block diagram, fig. 2.

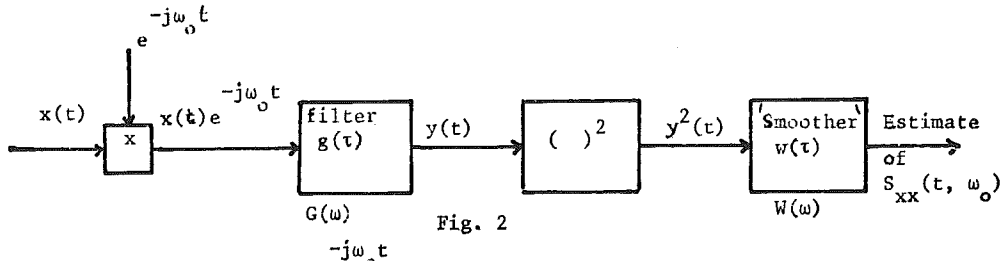


Fig. 2

The steps are (i) form $x(t)e^{-j\omega_o t} = y(t)$ (this complex demodulation procedure frequency shifts by ω_o); (ii) low pass filter $y(t)$; (iii) square the filtered signal; (iv) smooth the result. Whilst the technique is essentially the same as that used for stationary processes the important difference is the very careful choices to be made for the filters with weighting functions $g(\tau)$ and $w(\tau)$. Roughly speaking $g(\tau)$ must both 'low pass' $y(t)$ and yet also the modulating function $A(t, \omega)$, must be slowly varying compared with $g(\tau)$ so that the evolutionary trends are not lost. Similarly the filter $w(\tau)$ must smooth out sample fluctuations but the extent of the smoothing must be such as to achieve the variance reduction without losing the temporal information. This means there is a conflict in resolving temporal and frequency information. In fact one cannot obtain arbitrarily high resolution in both time and frequency together.

As noted before, there are other ways of estimating moving spectra which are essentially similar to that above, e.g., applying a sliding 'data window' to the record and Fourier transforming and then averaging (locally) the sequence of raw spectra obtained in this way. This is discussed in [8], which also contains examples of the estimation of the moving spectrum of aircraft noise measured during landing.

There are numerous applications of the evolutionary spectral density mainly to random vibration problems and a review paper by Van Marcke [9] contains a useful bibliography on this, and other, subjects.

2. NONLINEAR PHENOMENA

When a system is driven by a random excitation three problems can be addressed [9], namely, measurement, identification and prediction. More specifically, if we have measurements from processes involved we may wish to characterise the input or output signal in some way, e.g., estimate the probability distribution, or we may wish to identify the system based upon excitation and response measurements. Alternatively, we might model the system and excitation and attempt to predict the response. If the system is linear and the processes Gaussian then solutions to the problems are known. However, departures from linearity and Gaussianity mean that special methods must be brought to bear and the problems that arise have been treated in a number of ways and this is still a fertile research area. We shall only mention a few of the techniques that are available beginning with time domain methods and finishing with a frequency domain approach. A recent pair of survey articles by Roberts [10] and the paper by Van Marcke [9] (both addressed to random vibrations) are used as the basis of some of the general comments here.

Roberts lists and describes five approaches to the problem of predicting the response of nonlinear systems. They are (a) Markov methods; (b) equivalent linearisation methods; (c) perturbation methods; (d) functional series representations; and (e) simulation methods. We shall begin with the last, as it is conceptually the simplest. Sample functions of the excitation (either measured or synthetically generated) are used as the input to a mathematical model describing the system. Numerical methods are employed to produce corresponding sample response time histories. From these, summarising quantities may be generated, e.g., variances, higher moments, distributions, spectra, etc. A lot of data is needed to ensure that the summarising quantities themselves do not have a large spread. For stationary processes, time averaging is used to obtain the quantities but for nonstationary processes ensemble averaging over a *large number* of runs (see the example below) is required. Simulation also provides a valuable check for the validity of the analytic methods. Of the analytic methods listed above we shall briefly explain methods (a), (b) and (c).

A (vector) Markov process $\underline{x}$ is one which is characterised by its transition (conditional) probability density function $p(\underline{x}_1, t_2 | \underline{x}_2, t_1)$ and for a nonlinear dynamic system expressed in differential equation form $\dot{\underline{x}} = f(\underline{x}) + \underline{w}$ where $\underline{w}$ is the excitation, the state vector $\underline{x}$ is Markovian if $\underline{w}$ is a vector of white noise processes. It is possible to show further that the transition probability density function satisfies a partial differential equation of the diffusion type called the Fokker-Planck-Kolmogorov equation. Roberts [10] describes some cases for which exact solutions to the F-P-K equation have been given and also considers approximate solutions.

The equivalent linearisation method (b) is an approximate technique that is both popular and quite successful. The basic principle is that nonlinear systems are replaced by a set of 'equivalent' linear equations. This is sometimes referred to as the (stochastic) describing function technique. The essence of the method is that a nonlinear function $f(x)$ is approximated by the linear expression $n_o + n_1 x$. The quantities n_o, n_1 are obtained by minimising the mean squared error $E[(f(x) - (n_o + n_1 x))^2]$ with respect to n_o and n_1. The 'optimal' values n_o and n_1 require one to make an assumption about the distribution of x and often it is assumed that the equivalent linear system is Gaussian distributed. This technique, given above for a scalar quantity, can be extended to vector valued functions.

Applying this technique for example to a system characterised by

$$\dot{\underline{x}} = A\underline{x} + \underline{f}(x) + B\underline{w} \tag{9}$$

where $\underline{f}$ is a vector of nonlinearities and $\underline{w}$ is a vector of white noises, and if we assume symmetric nonlinearities (and so zero mean processes) an equivalent linear form for (9) is

$$\dot{\underline{x}} = (A + D)\underline{x} + B\underline{w}. \tag{10}$$

D is a matrix of equivalent linear gains and if $\underline{x}$ is assumed Gaussian D is a function of the zero lag correlations contained in matrix P $(= E[\underline{x}(t)\underline{x}^T(t)])$. It is then possible to write down an equation for $\dot{P}$ which is now nonlinear in P but which can be solved numerically.

As an example the vehicle on rough ground has been considered [4] with a nonlinear (square law) damper. Using fig. 1 again, where now we assume the damper is described by a square law nonlinearity we can write the equation as

$$\ddot{x} + \gamma \dot{x}|\dot{x}| + \omega_o^2 x = -\ddot{h}$$

where $x = (y - h)$ is the relative motion. Using a second order spatial model for the height profile and replacing the nonlinearity with an equivalent linear gain allows solution for the variances for accelerating vehicles. Figure 3 shows the propagation of $E[x^2]$ whilst the vehicle undergoes a constant acceleration from rest. We note that for low speeds the equivalent linearisation method is indistinguishable from the ensemble average of a 3000 run simulation. The computational saving thus offered is obvious since many runs are required for accurate estimates. (It is noted that the initial condition differs from that given in ref. [4] which was in error.)

Finally we mention technique (d), i.e., the functional series expansion. The literature on this method is extensive but a recent book by Shetzen [12] provides a comprehensive and readable account of the approach. The basis of the method is that the convolutional representation for linear systems is generalised for nonlinear systems. More specifically if $x(t)$ is input and $y(t)$ the output of a system, then the linear expression

$$y(t) = \int h(t_1)x(t - t_1)dt_1$$

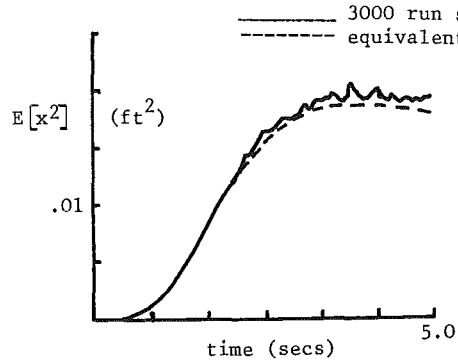


Fig. 3

is generalised to the form

$$y(t) = \sum_n \int\cdots\int_{n \text{ fold}} h_n(t_1, t_2, \ldots, t_n)x(t - t_1)x(t - t_2)\ldots x(t - t_n)dt_1\ldots dt_n$$

$$= \sum_n H_n[x(t)]. \qquad (11)$$

This is called a Volterra series and the $h_n(t_1, \ldots, t_n)$ are the Volterra kernels of the system. H_n, representing the n^{th} order convolution integral is called the n^{th} order Volterra operator. Whilst, as Shetzen shows, numerous interesting results may be developed for systems that admit such a description the Volterra series suffers limitations, namely, that estimation of the Volterra kernels is difficult and that the Volterra series (a generalised Taylor series) suffers convergence limitations. However, these disadvantages can be alleviated by using a set of orthogonal functionals suggested by Wiener formed from the Volterra functionals. These are called Wiener G-functionals because they possess the orthogonality property when the input to the system is Gaussian. This in turn means that the response of randomly driven nonlinear systems can be treated effectively using G-functionals. Furthermore, (and very significantly) generalised cross correlation techniques involving input and output allow determination of the higher order Wiener kernels of the Wiener G-functionals. For example, if the input $x(t)$ is white noise with spectral level A^2, the second order Wiener kernel is

$$\frac{1}{2A^2} E[\{y(t) - E[y(t)]\}x(t - \tau_1)x(t - \tau_2)].$$

Shetzen considers optimal (least squares) methods of nonlinear system identification using these methods for both white and nonwhite excitation and, since processes are assumed stationary, frequency domain versions of the results may be obtained.

The introduction of frequency domain methods now means that just as the correlation functions are generalised, so the spectra are generalised to higher order spectra (polyspectra) which are functions of more than one frequency variable. The first of these is the 'bispectrum', which is written

$$S_{xxx}(f_1, f_2) = \int_{-\infty}^{\infty}\int R_{xxx}(\tau_1, \tau_2)e^{-j2\pi f_1\tau_1 - j2\pi f_2\tau_2} d\tau_1 d\tau_2 \qquad (12)$$

where $R_{xxx}(\tau_1, \tau_2) = E[x(t)x(t + \tau_1)x(t + \tau_2)]$ is the generalised correlation function (or cumulant of order two). Inverting the equation and setting $\tau_1 = \tau_2 = 0$ gives

$$E[x^3(t)] = \int_{-\infty}^{\infty}\int S_{xxx}(f_1, f_2)df_1 df_2$$

thus showing that the bispectral density S_{xxx} is a decomposition over the f_1, f_2 plane of the third moment of $x(t)$, (in contrast to the ordinary spectrum which is a decomposition of the second moment of the process). For a Gaussian process, for example, $E[x^3(t)] = 0$ and so the bispectrum may be used as a measure of non-Gaussianity (and so perhaps nonlinearity) associated with a signal. In fact the bispectrum has been used for this purpose in the past. As far as published work on the bispectrum (and its higher order relatives) is concerned, Brillinger and Rosenblatt [13] have investigated the asymptotic theory of estimates of k^{th} order spectra and it has been applied to earthquakes [14], economic time series [15], and the study of sea waves [16]. Some recent work [17] has been concerned with statistical properties. The M.Sc. thesis by Perrochaud [11] describes the use of the 'cross bispectrum', i.e., $S_{xxy}(f_1, f_2)$, for input-output systems, to identify second order transfer functions of nonlinear systems. The thesis includes analysis of real data relating to human body vibrations and also analysis of a distorting loudspeaker.

3. CEPSTRAL ANALYSIS (HOMOMORPHIC SIGNAL PROCESSING)

In contrast to the previous discussions on nonstationary and nonlinear processes where the terms referred to the structure/generation of the processes, we now address a procedure that might be described as 'nonlinear processing' appropriate to a particular class of signals. We shall begin our description of cepstral analysis with reference to the analysis of 'echo-like' signals, i.e., signals which are composed of a 'wavelet' and one or more echoes which may overlap. The simplest form of this is shown in fig. 4, where the composite signal is $x(t) = s(t) + a_o s(t - t_o)$. The echo arrival time and the shape of the basic wavelet is usually obscured by noise, overlapping of echoes and wavelet, and distortion of of the echoes due to the characteristics of different transmission paths. Confronted with a signal of this type, the objective may be to estimate the echo arrival times, determine the wavelet shape or establish the 'transfer characteristics' of the transmission paths.

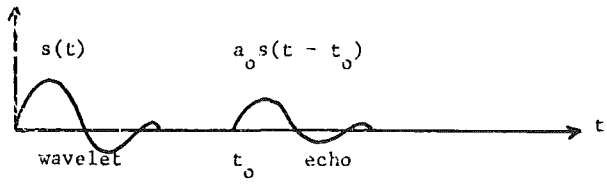


Fig. 4

This section will outline the use of cepstral analysis in the treatment of this problem. Much of the discussion and many of the examples are given in greater detail in the thesis by Peardon [18]. The references also include a readable account in the book by Childers and Durling [19] and the highly instructive pioneering paper by Bogert, Healy and Tukey [20]. It was in the latter paper that the term 'cepstrum' was first introduced to describe the result of forming the spectrum of a ($\log_e$) spectrum. In view of the 'interchanging' roles of time and frequency in this procedure a vocabulary was proposed to reduce likely confusion and some of these terms are defined later. Cepstral analysis was given further impetus with the publication by Oppenheim, Schafer and Stockham [21] which introduced the 'complex' cepstrum. In fact, the presentation in [21] proposed a general 'homomorphic system theory' dealing with signals combined nonlinearly, with the aim being to use appropriate transformations to render them in a form suitable for linear filtering techniques. We shall not consider this general theory explicitly but will deal only with its application to cepstral analysis.

3.1 The Power Cepstrum (Continuous Time)

The power cepstrum of a signal $x(t)$ was originally [20] defined as "the power spectrum of the natural logarithm of the power spectrum of the signal" and can be expressed mathematically as

$$x_p(t) = \left| \int_{-\infty}^{\infty} \ln |X(\omega)|^2 e^{-j\omega t} d\omega \right|^2 \qquad (13)$$

The flow chart in fig. 5 gives the sequence of operations needed to form $x_p(t)$.

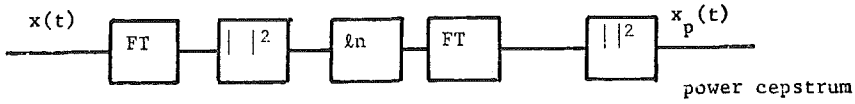


Fig. 5

It is appropriate to make the comment that written as (13) we should strictly refer to 'energy' rather than 'power' spectra. Various other definitions of the power cepstrum are used. The second Fourier transform is sometimes replaced by the inverse Fourier transform, and the final modulus and squaring operations are often omitted.

The use of this sequence of operations is clarified by considering the simple case of a wavelet and echo (fig. 4) where $x(t) = s(t) + a_o s(t - t_o)$ for which $|X(\omega)|^2 = |S(\omega)|^2(1 + a_o^2 + 2a_o \cos \omega t_o)$ so that

$$\ln |X(\omega)|^2 = \ln |S(\omega)|^2 + \ln(1 + a_o{}^2 + 2a_o \cos \omega t_o) \qquad (14)$$

demonstrating that the wavelet and echo information are *additively* combined after taking the logarithm. In fact the echo information is now a periodic function of frequency, and this periodicity can be investigated using usual Fourier techniques interpreting ω in the role of a 'time' variable. The fundamental 'frequency' of the periodic component depends on t_o but to emphasise the change of role this is referred to as a 'quefrency' so that the independent variable t of $x_p(t)$ (dimensionally time) is the 'quefrency axis'. The 'frequency series' $\ln |X(\omega)|^2$ can be filtered in the frequency domain (just as a time series) and such an operation is sometimes referred to as 'liftering', though it is probably far less likely to lead to confusion if the various 'data shaping' procedures that are employed are explained in relation to the context in which they are used.

Expanding the logarithm in powers of a_o and retaining only the first term yields

$$\ln |X(\omega)|^2 \simeq \ln |S(\omega)|^2 + 2a_o \cos \omega t_o .$$

The contribution of the echo to the log energy density is a cosine term which 'ripples' with *quefrency* t_o. The Fourier transform with respect to ω yields $F.T\{\ln |S(\omega)|^2\} + 2\pi a_o[\delta(t - t_o) + \delta(t + t_o)]$. (We shall omit the final modulus squaring operation in this discussion to avoid the introduction of cross product terms.) The echo has resulted in the appearance of 'spikes' at quefrencies $\pm t_o$. In practice, of course, the theoretical delta functions are replaced by functions dependent on the appropriate data window accounting for the finite operations involved.

The approximation of the logarithm has resulted in a single ripple characterising the echo. In fact, since the logarithm term is periodic a Fourier series can be employed, i.e., $\ln(1 + a_o{}^2 + 2a_o \cos \omega t_o) = A_o + \sum_{n=1}^{\infty} A_n \cos n\omega t_o$ where, if $|a_o| \leq 1$, the coefficients are $A_o = 0$ and $A_n = (-1)^{n+1}\, 2\, \frac{a_o^n}{n}$ in which case the Fourier transform of $\ln |X(\omega)|^2$ becomes

$$F.T\{\ln |S(\omega)|^2\} + \sum_{n=1}^{\infty} A_n \pi[\delta(nt_o - t) + \delta(nt_o + t)]$$

showing that the power cepstrum will have a sequence of spikes superimposed on the wavelet information rather than just one, owing to the nonlinear (spreading) effect of the logarithm, though the amplitudes of the spikes will decay as n increases.

Remarks

(1) The signal plus echo may be modelled as the convolution of $s(t)$ with a time function $\delta(t) + a_o\delta(t - t_o)$ and the operations in the power cepstrum analysis serve to 'separate' these convolved signals. In fact the above procedure can be regarded as a way of separating two arbitrary signals convolved with each other, e.g., if $x(t) = s(t) * h(t)$, then $\ell n|X(\omega)|^2 = \ell n|S(\omega)|^2 + \ell n|H(\omega)|^2$.

(2) Some important features of the cepstrum are apparent from this introductory discussion. We see that the wavelet should have a wide bandwidth, otherwise if $|S(\omega)|^2$ drops to a zero rapidly there is insufficient information about the ripple available. Furthermore, when the log of $|S(\omega)|^2$ is taken as the argument gets small the log spectrum usually has a very 'ragged' appearance which, owing to the great reduction of the dynamic range by the log operation (since $\ell n|X(\omega)|^2$ is now treated as the signal) has the effect of causing the cepstrum of the wavelet to be erratic. This 'noisy' behaviour makes it difficult to distinguish echo spikes from the wavelet cepstrum. Put another way, the log spectrum of the wavelet should be 'wide' and 'smooth' so that the wavelet cepstrum is 'low quefrency' allowing the echo cepstrum to be easily distinguishable from the wavelet cepstrum.

Discrete time

The formulation above does not account for the fact that computations will be done using discrete data. Indeed the majority of discussions of the cepstrum begin with the assumption that data is sampled. In fact the cepstrum does not exist for many continuous signals since if $S(\omega)$ goes to zero, then $\ell n|S(\omega)|^2$ is not defined, but for sampled signals the situation is easier to handle. The block diagram of fig. 6 is a satisfactory definition of the power cepstrum for sampled signals. The modulus squaring operation is omitted.

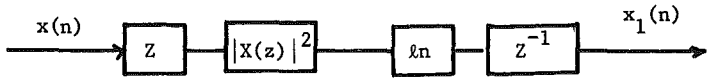


fig. 6

The equations are the appropriate discrete equivalents of those above. They are not developed in detail here since they are discussed in relation to the complex cepstrum later. Whilst the z-transform is indicated in fig. 6, computations are usually done using the Discrete Fourier Transform (FFT).

3.2 The Complex Cepstrum

The complex cepstrum of signal $x(t)$ is written $\hat{x}(t)$ and defined in fig. 7 (a), which is drawn for

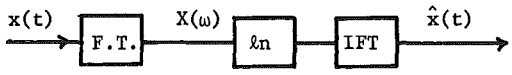


Fig. 7 (a)

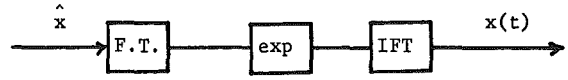


Fig. 7 (b)

continuous valued signals. It is as well to note that in spite of its name the complex cepstrum is *real valued* and the name has arisen to distinguish it from the previous power cepstrum since in this case a complex logarithm of $X(\omega)$ is used. It is this feature that makes complex cepstrum analysis more general than power cepstrum methods allowing the inverse operations to be performed to recover the original signal (fig. 7(b)). Postponing for the moment the definition of the complex logarithm it is instructive to compare this procedure with that of the power cepstrum for the single echo case, i.e. $x(t) = s(t) + a_o s(t - t_o)$. Following the operations in fig. 7(a) and approximating $\log(1 + a_o e^{-j\omega t_o})$ gives $\hat{x}(t) \simeq \hat{s}(t) + a_o\delta(t - t_o)$. This suggests immediately that the wavelet might be recovered if the 'spike' in $\hat{x}$ is 'removed' and the inverse operations of fig. 7(b) followed. This method does work but is complicated in practice by the following features. (a) Spikes usually occur at t_o and multiples of t_o due to the nonlinear logarithm operation. (b) The cepstrum of the wavelet may not be smooth enough to allow easy identification of echo spikes. (c) Noise on $x(t)$ causes severe degradation of $\hat{x}$, making it difficult to distinguish wavelet from echo information.

In spite of these difficulties, however, the cepstrum has come to occupy a place as a powerful way of deconvolving signals.

Discrete time formulation

The discrete time formulation is normally written using the z-transform even though the implementation is achieved using the FFT. This can be found in various sources (for example, [19, 22, 23]).

If $x_1(n)$ and $x_2(n)$ are two real sequences and if $x(n)$ is formed as the convolution of the two, i.e., $x(n) = x_1(n) * x_2(n)$, then the complex cepstrum of $x(n)$ is defined by $\hat{x}(n)$ (see fig. 8).

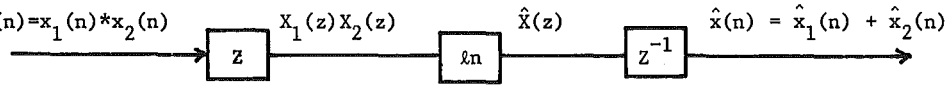


fig. 8

Evaluation is normally done on the unit circle $|z| = 1$ and so we substitute $z = e^{j\omega}$. Some comments are now in order regarding the definition of the complex logarithm. From fig. 8 we see

$$\hat{X}(z) = \log_e[X_1(z)X_2(z)] = \ell n|X_1(z)| + \ell n|X_2(z)| + j(\arg X_1(z) + \arg X_2(z)).$$

Whilst there is no ambiguity in defining the real part of the complex logarithm (when it exists), the imaginary part is multivalued (and so non-unique). The use of the principal value of the argument is inadmissible since (as we shall see) the resulting discontinuities are not permitted. These problems are circumvented in the following way. If $\hat{x}(n)$ is assumed to be a stable, real sequence then $\hat{X}(z)$ is analytic on $|z| = 1$ so that the real and imaginary parts $\hat{X}_R(e^{j\omega})$ and $\hat{X}_I(e^{j\omega})$ are continuous (and periodic) functions of ω, and since $\hat{X}_R(e^{j\omega}) = \ln|X(e^{j\omega})|$ and $\hat{X}_I(e^{j\omega}) = \arg(X(e^{j\omega}))$ then these functions too, are continuous (and periodic). So it is the following conditions that dictate how the imaginary part of the complex logarithm should be defined, i.e., $\arg X(e^{j\omega})$ should be (i) uniquely defined; (ii) continuous; (iii) periodic with period 2π; (iv) the value at $\omega = \pi$ is zero. These aspects are commented upon in the 'phase unwrapping' discussion under computational considerations.

Once the phase is satisfactorily defined then as we have seen above, the complex cepstrum separates convolved signals into additive form, where they may be 'liftered' and then the inverse operations performed to recover one or other of the signals.

3.3 Multi-echo Signals, Distortional Channels and Noise

In this section we point out aspects of the application of cepstral methods to problems other than those restricted to a single undistorted echo.

Multi-echo signals

In many situations there may be several echoes present in a signal (e.g., reflection seismology) and even when the echoes are scaled replicas of the signal (i.e., undistorted) the analysis can become extremely involved, but we shall make some observations that are useful in practical situations.

An appropriate model (written in continuous time) for a composite signal comprising a basic wavelet and N echoes is

$$x(t) = s(t) + \sum_{n=0}^{N-1} a_n s(t - t_n)$$

so that

$$\ln X(\omega) = \ln S(\omega) + \ln\left[1 + \sum_{n=0}^{N-1} a_n e^{-j\omega t_n}\right]$$

Assuming the log expansion for the echo term to be valid the problems become apparent since not only does each echo location contribute a spike, so do their multiples and sums of combinations of these, thus complicating the appearance (and hence interpretation) of the cepstrum.

A possible remedy is 'windowing' to limit the number of echoes under consideration. A particularly effective window is merely a decaying exponential applied from $t = 0$ which has the effect of simply weighting (reducing) the magnitude of the 'spikes' obtained. (Windows are useful too for data that is nonstationary in character, e.g., speech [22] and seismic data [23].)

Distortional channels

The echoes considered so far have been scaled replicas of the basic wavelet. In many practical situations, however, the echoes will be distorted due to influences along the paths (channels) along which they travel. The recorded signal may be modelled as the source wavelet convolved with the channel(s) impulse response. Such filtering has the effect of spreading the original pulse-like wavelet making echo detection and wavelet recovery more difficult.

Another obvious distortional effect is that of additive interference (noise). This can, of course, arise from a number of sources and is the major degrading influence in the effectiveness of cepstral methods.

Convolutional distortion

Consider a measured signal made up of a wavelet plus a distorted echo. This can be modelled as (fig. 9) $x(t) = s(t) + y(t - t_o)$ where $y(t) = h(t) * s(t)$. The function $h(t)$ represents the impulse response function of the channel.

Hassab and Boucher [24] have considered this in detail and we outline the treatment here. It follows from the model that

$$\ln X(\omega) = \ln S(\omega) + \ln\left[1 + H(\omega)e^{-j\omega t_o}\right].$$

Assuming the logarithm expansion is valid results in the complex cepstrum

$$\hat{x} = \hat{s} + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} h_n(t - nt_o) \qquad (15)$$

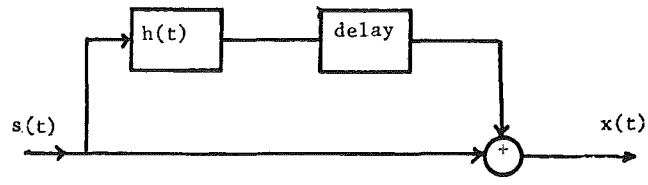


Fig. 9

where $h_1(t) = h(t)$ and $h_{n+1}(t) = h_n(t) * h(t)$, $n = 1, 2....$ Equation (15) shows that the impulse response function occurs in the complex cepstrum as a series of n-fold auto-convolutions at integer multiples of the time delay. In fact, provided severe overlap does not occur in the region near t_o, it is possible to obtain an estimate of $h(t)$ directly from the cepstrum [25].

Additive distortion

A model for additive distortion on a single echo is $x(t) = s(t) + a_o s(t - t_o) + y(t)$ where $y(t)$ is 'noise'. Analysis can be carried out (ref. [18]) formally for this case but the algebra is tedious. Furthermore, without any particular structure to $y(t)$ it is only possible to predict that functional shapes involving $Y(\omega)$ and $S(\omega)$ will degrade the cepstra. In fact, wide band noise very soon degrades cepstra to the point of not being able to pick out the underlying features. Usually some

form of signal averaging is required to deal with this situation either in the time domain before cepstral analysis begins or in the cepstral domain (or equivalently in the frequency domain, e.g., averaging the $\ln|X(\omega)|^2$ function for power cepstra).

3.4 Computational Considerations

This section briefly refers to three aspects that arise in the computation of cepstra that call for particular care if the method is to be effective.

(a) Cepstrum aliasing

The problems associated with sampling continuous data have been discussed elsewhere and an adequate sampling rate alleviates aliasing. Similar precautions may be necessary to avoid cepstrum domain aliasing, which can occur if the ripple in the frequency domain is too quick. There are two possible remedies for this problem. (i) *Appending zeros* to the original time history ensures a finer sampling of the complex logarithm in the frequency domain and an equivalent extension of the cepstrum domain and the 'aliasing' problem is reduced. (ii) *Exponential weighting* of the original time history reduces the amplitudes of the high quefrency harmonics and so less overlapping occurs.

(b) Phase unwrapping

We have already seen that the complex logarithm needs careful definition to ensure that the complex cepstrum is 'properly' defined, and we noted conditions that the imaginary part of the logarithm (i.e., $\arg X(e^{j\omega})$) should satisfy. The procedure ensuring that $\arg X(e^{j\omega})$ has these properties has come to be called 'phase unwrapping'. Various methods may be used but Tribolet [26] has developed an approach which is widely used. It involves an iterative scheme which ensures a suitably fine resolution in the frequency domain (where required) in order to 'follow' the phase variation.

(c) Liftering

Various forms of liftering may be employed and the thesis by Peardon [18] describes some of them.

3.5 Applications

Cepstral methods have been applied widely and a short list of some applications follows: speech signal decomposition [22], seismic signal processing [23,18], room impulse response studies [29], the measurement of reflection coefficients [25], ground reflection assessment [27] and pattern recognition [28].

References

1. J.S. Bendat and A.G. Piersol. Random Data: Analysis and Measurement Procedures. Wiley Interscience 1971.
2. M.B. Priestley and T.S. Rao. A test for non stationarity of time series. J. Roy. Stat. Soc. B. 31, 1. 1969.
3. J.K. Hammond and R.F. Harrison. Nonstationary response of vehicles on rough ground - a state space approach. Proc. of the Winter Annual Meeting ASME, Chicago. 1981.
4. R.F. Harrison and J.K. Hammond. The response of vehicles on rough ground. Proc. of the Spring Meeting of the Inst. of Acoustics. Newcastle, England. 1981.
5. M.B. Priestley. Evolutional spectra and non-stationary processes. J. Roy. Stat. Soc. B.27, 1965.
6. M.B. Priestley and H. Tong. On the analysis of bi-variate non-stationary processes. J. Roy. Stat. Soc. B.35, 1973.
7. J.K. Hammond. Evolutional spectra in random vibrations. J. Roy. Stat. Soc. B. 35, 1973.
8. Y.H. Tsao. A significance test for nonstationarity of random processes. ISVR Memorandum 516, Southampton University. 1981.
9. E.H. Van Marcke. Some recent developments in random vibration. Applied Mechanics Reviews, Vol. 32, No. 10. 1979.
10. J.B. Roberts. Response of nonlinear mechanical systems to random excitations. Shock and Vibration Digest Part I, Vol. 13, No. 4. April 1981. Part 2, Vol. 13, No. 5, May 1981.
11. J. Perrochaud. Bispectral analysis of nonlinear systems. M.Sc. Thesis, ISVR, University of Southampton. 1981.
12. M. Shetzen. The Volterra and Wiener Theories of Nonlinear Systems. Wiley-Interscience. 1980.
13. D.R. Brillinger and M. Rosenblatt. Computation and interpretation of k^{th} order spectra. Proc. of an Advanced Seminar on Spectral Analysis of Time Series. Ed. R. Harris. Wiley. 1967.
14. R.A. Haubrich. Earth noise 5 to 500 millicycles per second. Jnl. of Geophysical Research, Vol. 70, No. 6, March 1965.
15. M. Godfrey. An exploratory study of the bispectrum of economic time series. Appl. Stat. 14, 1965.
16. K. Hasselmann, W. Munk and G. MacDonald. Bispectra of ocean waves. Time Series Analysis. Ed. M. Rosenblatt. J. Wiley, 1963.
17. M. Gabr and T.S. Rao. A note on the estimation of the bispectral density function of a stationary time series. Tech. Report 123, Sept. Department of Mathematics, UMIST, Manchester. 1979.
18. L.G. Peardon. Aspects of cepstral analysis and epoch detection. M. Phil. Thesis. Dept. of Maths. Portsmouth Polytechnic. 1979.
19. D.G. Childers and A. Durling. Digital Filtering and Signal Processing. West Publishing Co. 1975.
20. B.P. Bogert, W.J.R. Healy and J.W. Tukey. The quefrency alanysis of time series for echoes: cepstrum, pseudo autocovariance, cross cepstrum and saphe cracking. Time Series Analysis. Ed.

M. Rosenblatt, Wiley. 1963.

21. A.V. Oppenheim, R.W. Schafer and T.G. Stockham, Jr. Nonlinear filtering of multiplied and convolved signals. Proc. IEEE, Vol. 56, No. 8. Aug. 1968.

22. A.V. Oppenheim and R.W. Schafer. Digital Signal Processing. Prentice Hall, Inc. 1975.

23. J.M. Tribolet. Seismic applications of homomorphic signal processing. Prentice Hall, Inc. 1979.

24. J.C. Hassab and R. Boucher. Analysis of signal extraction, echo detection and removal by complex cepstrum in presence of distortion and noise. Journal of Sound and Vibration, Vol. 40, No. 3, 1975.

25. J.S. Bolton and E. Gold. The measurement of reflection coefficients in-situ using cepstral techniques: simulations. Proc. Autumn Conference of Inst. of Acoustics, Windermere. Nov. 1980.

26. J.M. Tribolet. A new phase unwrapping algorithm. IEEE Trans. on Acoust., Speech and Sig. Proc., Vol. ASSP-25, No. 2, 1977.

27. A.A. Syed, J. Brown, M. Oliver and S.A. Hills. The cepstrum: a viable method for the removal of ground reflections. Journal of Sound and Vibration, Vol. 71, No. 2. July 1980.

28. D.W. Thomas. Vehicle Sounds and Recognition. Pattern Recognition - Ideas in Practice. Ed. B. Batchelor. Plenum Press. 1978.

29. J.K. Hammond and J. Mourjopoulos. Cepstral methods applied to the analysis of room impulse response. Proc. Inst. of Acoustics Autumn Conference Windermere. Nov. 1980.

PARAMETRIC METHODS IN SIGNAL ANALYSIS
(with particular reference to autoregressive spectral estimation)

J.K. Hammond
Institute of Sound and Vibration Research
University of Southampton
Southampton SO9 5NH
England

SUMMARY

Data may be analysed using parametric or nonparametric methods, and often it is not clear which is the most appropriate technique. In this chapter we describe one particular aspect, namely, parametric spectral analysis using the Maximum Entropy Method. The treatment is tutorial in nature, describing general linear models for time series and how the Maximum Entropy Spectral Density relates to autoregressive processes. A list of applications of the method is included. The chapter concludes by noting a few other parametric approaches to data analysis and system characterisation.

1. INTRODUCTION

Data may be examined and described without referring directly to the process by which it was generated, though interpretation of the results requires information about the mechanism of data generation. Conventional spectral analysis is an example of a data analysis technique that does not rely on such knowledge and is referred to as a nonparametric procedure [27]. However, if prior knowledge of the process is available then it is often appropriate to build this information into the data processing/reduction techniques. This could mean that a reasonably detailed mathematical model describing the phenomenon is constructed, which is characterised by a restricted set of parameters, and 'data analysis' is then concerned with estimation of these parameters from measured data. For example, some vibrating systems are adequately described by coupled linear ordinary differential equations and it is logical to 'fit' suitable mass, stiffness and damping parameters to a chosen model so as to best account for some measured data.

Even when a well defined prior mathematical model is not available it is sometimes attractive to try and fit classes of model (e.g., linear difference equation forms) to data, since they may give insight into the problem that was previously obscured. It is obvious though that one should guard against the tendency to constrain the analysis to fit data to that which appears to be (currently) the most suitable model. Often then, the choice between a parametric or nonparametric approach for a specific application is not obvious.

The articles [28, 6] discuss several aspects of the points noted above. It is sufficient to say that parametric methods are well established with wide application. The literature on the subject is vast and in this chapter we will describe only one particular aspect, namely, parametric spectral analysis using the Maximum Entropy method.

At the end of the chapter just a few other parametric approaches to data analysis and system characterisation are noted very briefly.

2. MAXIMUM ENTROPY SPECTRAL ANALYSIS

Conventional methods of spectral estimation use filtering methods, windowed correlation functions or 'segment averaging'. Just over a decade ago new procedures were introduced that offered the possibility of improved resolution over the established methods. In this chapter we outline the basis of the approach invented by Burg [1,2,3] and generally referred to as the Maximum Entropy method (abbreviated MEM). Many papers have been written on the subject, some of which have been collected together in a useful volume edited by Childers [3], and the material for this chapter is drawn from this and other sources listed in the bibliography. It is appropriate to begin this tutorial treatment by considering the structure of a time series. In common with all the texts on this subject the processes involved are assumed to be discrete in time.

2.1 Time Series Models [4,5]

It seems self-evident that the procedure used to estimate a spectrum of a time series should depend upon the method in which the data is generated and so if the mechanism of data generation is unknown then there is no one 'correct' method of spectral estimation. A general time series x_n will have some complex structure reflected in its mean, covariance, etc., and various models have been proposed to account for the behaviour. A common approach is to regard x_n as having arisen from appropriate operations on white noise w_n (defined by $E(w_n w_k) = 0$, $n \neq k$ and $E(w_n^2) = \sigma^2$). We shall restrict our discussions to *linear* operations and *stationary* time series.

A general linear model for x_n can be written [5] as $\sum_{k=-\infty}^{\infty} g_k x_{n-k} = w_n$. When expressed in this form it is perhaps not apparent that w_n is an 'input' and x_n an 'output'. Rather, the reason for choosing this structure for the model is that we are trying to find some relationship between past, present and future values of the time series x_n which reduces it to white noise (having elementary properties). The model as posed is acausal (anticipative) and to ensure that x_n depends only on past values we set $g_k = 0$ for $k < 0$ so that the model becomes

$$\sum_{k=0}^{\infty} g_k x_{n-k} = w_n. \qquad (1)$$

Assuming that we can 'invert' this equation we write

$$x_n = \sum_{k=0}^{\infty} h_k w_{n-k} \qquad (2)$$

This is an alternative form for the general acausal linear model where x_n is regarded as a sum of

weighted values of the white noise. It is in this form that h_k can be regarded as a system impulse response sequence, w_n a white noise input sequence and x_n the output sequence.

AR, MA and ARMA models

The models (1) and (2) are characterised by the *infinite* set of parameters g_k or h_k. The calculation of these parameters by fitting such models to observed data would provide the required spectrum. The spectrum of x_n following from (2) is $|H(z)|^2\sigma^2$ where

$$H(z) = \sum_{k=0}^{\infty} h_k z^{-k} \quad \text{and} \quad z = e^{+j\omega} .$$

However, for practical situations it is necessary to reduce the models to those containing only a *finite* number of parameters.

A finite version of (1) is (with $g_o = 1$)

$$x_n + g_1 x_{n-1} + g_2 x_{n-2} + \ldots + g_p x_{n-p} = w_n. \tag{3}$$

This is called an autoregressive (AR) model of order p.

A finite version of (2) is (with $h_o = 1$)

$$x_n = w_n + h_1 w_{n-1} + h_2 w_{n-2} + \ldots + h_q w_{n-q}. \tag{4}$$

This is called a moving average (MA) model of order q.

Rather than attempting to approximate a general linear model by either an AR or MA model of sufficiently high order we can use a mixed form to obtain a better fit with fewer parameters, namely the autoregressive - moving average (ARMA) form. (N.B. We shall now change our notation, having settled on this 'finite' form and write the auto-regressive parameters as a_i and the moving average parameters as b_i.)

An ARMA (p, q) (mixed autoregressive/moving average model of order (p, q) is

$$x_n + a_1 x_{n-1} + \ldots + a_p x_{n-p} = b_o w_n + b_1 w_{n-1} + \ldots + b_q w_{n-q} \tag{5}$$

(often b_o is set to unity without any loss of generality).

This transfer function for this system is

$$H(z) = \frac{b_o + b_1 z^{-1} + \ldots + b_q z^{-q}}{1 + a_1 z^{-1} + \ldots + a_p z^{-p}}$$

and the power spectral density of x_n is $|H(z)|^2\sigma^2$ with $z = e^{j\omega}$.

As noted at the beginning of this section, if we wish to estimate the power spectral density from a finite length of data it is only sensible to choose an estimation procedure appropriate to the data model, i.e., we must know whether the data is best described by AR, MA or ARMA processes, though in a practical situation we may have few clues as to which model is most appropriate. The most general of the three, the ARMA model, has recently been the focus of much attention (e.g., [5]). The MA model can be treated by the 'classical' windowed correlation method [4], and the AR model arises naturally out of the maximum entropy method.

2.2 Maximum Entropy Spectral Estimation

As noted above, the 'conventional' approach to spectral estimation matches the finite MA model since the autocovariance function is of finite length and so windowing effects might be discounted. But if the MA model is of infinite order (equation (2)) then so is the autocovariance function and so windowing results in an information loss resulting in bias/resolution problems. If such an infinite order MA process could be described by a finite order AR or ARMA process this might be a useful model on which to base the estimation procedure. We might consider some method of estimating the (say) AR parameters directly or 'suitably' extending the autocovariance function in some logical manner rather than arbitrarily truncating it after 'running out of data'. Interestingly, it turns out that these procedures are equivalent (for AR processes) and constitute the ME method. However, we have yet to account for why the concept of 'entropy' arises in the description, and this we now do rather briefly. (The references [7,8,1] and particularly [9] are illuminating and serve as the basis of this section.)

The problem of spectral analysis is one of data reduction and analysis and it seems self-evident that a basic principle of data reduction should be that *any transformation imposed on experimental data shall incorporate and be consistent with all relevant data and be maximally noncommital with regard to unavailable data.* The act of setting estimates of the autocorrelation function to zero outside a particular range is hardly being noncommital. It would seem more appropriate to extend the autocovariance function in some 'reasonable' way, but whilst still being *noncommital* with regard to unavailable data. It is appropriate to introduce the concept of entropy at this stage. The 'unavailable data' referred to is a lack of *information*, and information and entropy are related. In fact, *entropy* is a measure of uncertainty, or disorder of a system. To see how the notion of maximising entropy enters the picture we attempt to state the basic principle of data reduction in rather more specific terms.

Suppose we possess some data about a signal, x_n, in the form of values of the autocovariance function, i.e., $R(0)$, $R(\pm 1)$... $R(\pm p)$ and we wish to produce from this *finite length* sequence a spectral density of a signal in the range $-\frac{1}{2} < f < \frac{1}{2}$. We know that the autocovariance sequence $R(n)$ and spectral density $S(f)$ are related by

$$R(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S(f)e^{+j2\pi fn}df \qquad (6)$$

but this does not provide a *unique* way of obtaining an estimate of S from R. The traditional way simply chooses one lag window from an infinite set in the inversion, thus contravening the basic principle of data reduction by assuming $R(n)$ is zero for $|n| > p$ and, further, 'rejecting' good data in any tapering process. Since we should be noncommital with regard to the unavailable covariance function estimates (or our lack of information about them) it is desirable to be able to quantify our *ignorance of the original signal* as expressed by the spectral density $S(f)$ and maximise this, but still taking full account of the measured data as constraints. What we still need, however, is an expression for the ignorance (entropy) given in terms of the spectrum $S(f)$. The usual definition of entropy (strictly the relative entropy) given by

$$H = -\int_{-\infty}^{\infty} p(u)\ \log\ p(u)du,$$

where $p(u)$ is the signal amplitude probability density function is not useful but Bartlett (see references [7, 9]) has argued that the (Gaussian) signal with spectral density $S(f)$ can be regarded as having been created by passing white noise through a linear filter, and the *entropy difference* between the input process (which has the maximum entropy of all processes with the same variance) and the output is the entropy loss ΔE where

$$\Delta E = -\int_{-\frac{1}{2}}^{\frac{1}{2}} \log\ (S(f))df. \qquad (7)$$

From (6) the equation (7) is written

$$\Delta E = -\int_{-\frac{1}{2}}^{\frac{1}{2}} \log\left[\sum_{n=-\infty}^{\infty} R(n)e^{-j2\pi fn}\right]df \qquad (8)$$

Now minimising ΔE (or maximising the entropy) with respect to the unknown $R(n)$, i.e., $|n| > p$, subject to the constraint that the spectrum is consistent with the known $R(n)$, i.e., $|n| < p$ leads to the maximum entropy spectrum and is the spectral estimate that arises from making the least number of assumptions about the unmeasured data, whilst remaining consistent with the known information.

Thus the maximum entropy method finally reduces to a calculus of variations problem, solved using Lagrange multipliers and the resulting MEM spectrum is [8] of the form

$$\hat{S}_{MEM} = \frac{K}{\left|1 + \sum_{k=1}^{p} \alpha_k e^{-j2\pi fk}\right|^2} \qquad (9)$$

where K is a constant.

Quite clearly the spectral estimate has a form similar to the transfer function of an all-pole digital filter, being in fact the squared magnitude of the frequency response of the all-pole filter excited by white noise.

2.3 MEM and AR Processes

As noted above, the MEM spectrum is equivalent to the spectrum of a white excited AR filter. This correspondence between MEM and AR processes was formalised by VandenBos [10] for Gaussian processes where the starting point for the AR form is

$$x_n = -a_1x_{n-1} - a_2x_{n-2} - \ldots - a_px_{n-1} + w_n \qquad (10)$$

and so determining the extension to the autocovariance function is equivalent to calculating the parameters a_i of the AR model (10). Indeed, knowledge of the coefficients and $R(n)$ for $n = 0, \ldots, p$ implies knowledge of $R(n)$ for all n if x_n is described by (10). This follows from (10) by multiplying by x_{n-k} for $k > 0$ and taking expectations. This yields

$$R(k) = -a_1\ R(k - 1) - a_2\ R(k - 2) \ldots - a_pR(k - p), \quad k > 0 \qquad (11)$$

Clearly, knowledge of $R(0), \ldots, R(p - 1)$ and the a_i is sufficient to go on generating $R(n)$ for $n \geq p$. We note in passing that when $k = 0$ we can show

$$R(0) = -a_1R(1) - a_2R(2) + \ldots - a_pR(p) + \sigma^2. \qquad (12)$$

Of course, in the estimation problem we only know the autocovariance function estimates and not the a_i, which must be determined to calculate the spectrum.

Before considering the problem of estimation of the a_i it is worth noting the interpretation of the AR process (10) as one that identifies x_n with a value that is predicted from the previous values of the process with an error $e_n = w_n$ and the coefficients $a_1, \ldots, a_p$ define a p point prediction filter with $w_n = x_n + a_1x_{n-1} \ldots + a_px_{n-p}$, so the *prediction error* filter has coefficients $(1, a_1, \ldots a_p)$. (See reference [13].)

2.4 Estimation of AR Coefficients

We shall now consider how the MEM spectrum may be computed using (9). Clearly it is necessary to determine (i) the length p of the prediction filter and then (ii) the coefficients a_i. However, we shall see in section 2.5 that the choice of p requires a knowledge of the a_i and so we shall assume here that p is fixed and, subject to that, calculate the estimates $\hat{a}_i$ of a_i.

We emphasise that the discussions so far have related to the definition of the MEM spectrum using exact values of the autocovariance function. No mention has been made of the use of *estimates* of the autocovariance function to produce estimates of the MEM spectra. The estimates of the autocovariance function contain the effects of additive noise and data truncation. The first problem is considered by Ables [9] who suggested the modification that the autocorrelation be constrained to pass close to measured values. The problem of data truncation amounts to a violation of the basic principle of data reduction and this is discussed again in the next section.

Coefficient estimations

(a) The Yule Walker (YW) Method.

A widely used method of obtaining estimates $\hat{a}_i$ of a_i follows directly from (11) which can be written out in matrix form for $k = 1,\ldots, p$ as

$$\begin{bmatrix} R(0) & R(1) & R(2) & \ldots & R(p-1) \\ R(1) & R(0) & R(1) & \ldots & R(p-2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ R(p-1) & R(p-2) & R(p-3) & \ldots & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_p \end{bmatrix} = \begin{bmatrix} -R(1) \\ -R(2) \\ \cdot \\ \cdot \\ -R(p) \end{bmatrix} \qquad (13)$$

From the above we note that knowledge of $R(0), \ldots, R(p)$ gives enough information to obtain the a_i. Any usual (e.g., Gauss elimination) technique may be used to solve (13) but note that the coefficient matrix is symmetric and has a diagonal symmetry of the coefficients (i.e., elements on leading diagonals are equal). Such matrices are called Toeplitz matrices and their special structure can be used to speed the solution using a recursive procedure originated by Levinson and improved by Durbin (see [8]).

To write down the spectral estimate (9) requires not just the coefficients a_i but also the scale factor σ^2, and this follows from (12) which when this final equation is added to the set (13) results in the alternate form

$$\begin{bmatrix} R(0) & R(1) & \ldots & R(p) \\ R(1) & R(0) & \ldots & R(p-1) \\ \cdot & \cdot & & \cdot \\ R(p) & R(p-1) & & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \cdot \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \cdot \\ 0 \end{bmatrix} \qquad (14)$$

As noted in [8], the correspondence between the AR process and the prediction of x_n from past data means that σ^2 may be identified with the prediction error variance E_p. Claerbout [12] gives a recursive solution to equation (14). We emphasise that this formulation has used 'theoretical' values for $R(n)$. In any practical estimation problem we only have the measured data x_n from which we might make estimates $\hat{R}(n)$ and from the $(p + 1)$ equations (14) we solve for the $(p + 1)$ unknowns $a_1, a_2, \ldots a_p$ and σ^2.

The *Yule Walker* (YW) estimates of the AR coefficients may be criticised since (i) the AR coefficients should be estimated so as to be noncommital with regard to unavailable information. However the usual methods of estimation of $\hat{R}$ imply $\hat{R} = 0$ outside the available data length, thus contradicting the principle of maximum entropy. (ii) If $\hat{R}$ is estimated using the unbiased estimator it is sometimes possible to obtain estimates that do not constitute a valid autocovariance matrix [12]. (iii) The YW estimates are sensitive to rounding errors [8] when poles of the linear system lie near $|z| = 1$.

(b) The Burg method of coefficient estimation [1,2,12]

A method for the computation of the $\hat{a}_i$ proposed by Burg avoids the problems (i) and (ii) above. The need for calculation of the autocovariance of the data is sidestepped altogether and the scheme essentially uses the interpretation of the MEM spectrum as the spectrum of the process which is the output of an AR filter. In fact the method obtains the prediction error filter coefficients directly using least squares methods. We follow Claerbout's [12] explanation in what follows.

Suppose we have a data segment $x_0, x_1, \ldots, x_N$. Then a simple (two point) prediction filter for x_n is $\tilde{x}_n = -a\, x_{n-1}$ and the prediction error is $e_n = (x_n + a\, x_{n-1})$. The parameter is obtained by finding a that minimises

$$\sum_{n=1}^{N} e_n^2.$$

This straightforward approach, however, suffers the disadvantage that the resulting a may not satisfy $|a| < 1$ and so the prediction error filter is non-minimum phase.

Burg noted that for stationary random signals a prediction filter works just as well when reversed and run backward in time, i.e., $x_{n-1} = -a\, x_n$ (prediction filters defined in section 2.3 depend only on the autocovariance of the data and not the data itself and so the same prediction filter arises from the time series and a time reversed version). This being so, the average error observed when this two point filter is operated in reverse is averaged with the forward error, i.e.,

$$E(a) = \frac{1}{2(N-1)} \sum_{n=1}^{N} \{(x_n + a\, x_{n-1})^2 + (x_{n-1} + a\, x_n)^2\} \qquad (15)$$

Two important aspects/results of this formulation are noted. (i) Minimisation of E with respect to a always leads to a value $|a| < 1$ (i.e., a minimum phase prediction error filter). (ii) The filter is *not* run off the ends of the data sample and so no assumptions about the time series outside the sample interval are required.

This sample scheme for a two point filter seems easily generalised to a three point filter, i.e.,

$$E(a_1, a_2) = \frac{1}{2(N-2)} \sum_{n=2}^{N} \{(x_n + a_1 x_{n-1} + a_2 x_{n-2})^2 + (x_{n-2} + a_1 x_{n-1} + a_2 x_n)^2\} \qquad (16)$$

and simply minimising this with respect to a_1 and a_2 is required. Unfortunately it is possible that the filter weights that arise do not constitute a prediction error filter, i.e., the Toeplitz matrix corresponding to the estimated filter is not non-negative definite. Also the estimate of $R(1)$ from the three point filter does not agree with that estimated by the two point filter. Burg resolved this problem by adopting a recursive technique relating the three point filter to the two point filter. In fact, this recursion is very similar to that used to solve the Yule-Walker equations.

The essence of the procedure is that the coefficient vector of order 3 can be built up from that of order 2 by

$$\begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ a \\ 0 \end{bmatrix} - c \begin{bmatrix} 0 \\ a \\ 1 \end{bmatrix} \qquad (17)$$

and instead of trying to calculate a_1 and a_2 by minimising (16) directly, a is found first by minimising (15) and then parameter c (in (17) is found by minimising

$$E(c) = \frac{1}{2(N-2)} \sum_{n=2}^{N} \{(x_n + a(1-c)x_{n-1} - cx_{n-2})^2 + (x_{n-2} + a(1-c)x_{n-1} - cx_{n-2})^2\}$$

The parameter a remains fixed and minimisation if (18) with respect to c results in a minimum phase three point prediction error filter. This procedure can be extended to higher order processes and Fortran programs are available in Ulrych and Bishop's paper [8] for both Yule-Walker and Burg estimates.

To complete the requirements for the calculation of the spectral density the innovation (white noise excitation) variance estimate is calculated as with the Yule-Walker method.

2.5 Choice of Filter Length

The MEM spectral density estimate requires the length p (order of the AR process) to be known. Without prior information this is a difficult task. Underestimates of p tend to lead to 'oversmoothed' spectra and overestimates should be avoided since (a) computation should be minimised; (b) ill-conditioning of the equations increases with numbers of poles; (c) spurious ripples occur in the spectrum.

Several authors have considered the problem of order determination and the most successful and widely used method of order determination can be attributed to Akaike [16, 25] (though there are others [23]). The criteria are based upon the use of the 'error variance' which is monitored at each iteration. Akaike's approach is to combine into a single function both the error and the number of parameters, which should demonstrate a minimum beyond which point an increase in order is not appropriate. Three such functions are given below; the first two being due to Akaike and the third to Parzen. The terms FPE, AIC and CAT stand for final prediction error, Akaike's information criterion and criterion autoregressive transfer function, respectively. (E_p denotes the prediction error variance.)

$$FPE_{(p)} = \frac{N + p + 1}{N - p - 1} E_p$$

$$AIC_{(p)} = N \log E_p + 2p$$

$$CAT_{(p)} = \left(\frac{1}{N} \sum_{i}^{p} \frac{1}{E}\right) - \frac{1}{E_p}$$

An investigation of all three criteria undertaken by Beamish and Priestley [15] concluded that all three formulations lead to similar results except for 'awkward' time series, i.e., modelled by MA terms. See also the work of Tong [11], who has discussed the use of AIC for order determination.

2.6 Properties of MEM Spectra

Variability

Ulrych and Bishop [8] report on the work of Kromer who determined the asymptotic properties of the MEM estimator which are: (i) the estimator is asymptotically unbiased and (ii) the variance of the estimator is $(2/n)S^2(f)$ where $n = N/p$. This holds when N and p are large and $S(f)$ is smooth.

Akaike [16] has noted that confidence intervals can be evaluated for MEM spectra if the model order is assumed to be correct and Kromer's conclusion accepted. Baggeroer [17] has discussed the problem of obtaining confidence intervals more closely.

Resolution

An often noted reason for using MEM estimates is that this is a high resolution approach that is not restricted to the finest resolution being the reciprocal of the data length and numerous authors have attempted to quantify this.

Marple [18] gives an approximate result for the 'two sinusoid' frequency resolution δf, which depends on the signal-to-noise ratio (SNR) which for SNR > -3 dB is $\delta f \simeq 1.03/p(\Delta t)\ SNR(p+1)^{.31}$ (Δt = 1/sample rate) and simulation studies by Dyson and Rao [19] confirm this. Durrani and Taylor [20] describe the side lobe structure for data containing pure tones in wide band noise.

2.7 Applications

MEM spectral analysis has been applied to a wide range of topics, including speech, geophysics, astronomy, EEG analysis (references [3,13]), spatial array data [33], etc. As far as structural vibrations are concerned the work reported in references [14 and 21] are particularly relevant. The work of Vandiver and Campbell [21] is concerned with applying MEM spectral analysis to offshore structures where short data records may have to be used owing to the nonstationary nature of the wave excitation. Laight's work [14] is a detailed study of the use of MEM on vibrating systems of varying complexity (including a model offshore rig). The detailed conclusions are too lengthy to include here but the general conclusions are that lightly damped systems (with separated modes) are suitable candidates for AR modelling, and since peaks in spectra are well represented (with 'correct' choice of model order) then damping ratios can be estimated.

3. OTHER METHODS

We shall refer in very brief terms to a few other paramtric methods. The *Prony Series* is a method of characterising signals in the form $x(t) = \sum c_j e^{s_j t}$ where the c_j, s_j may be complex. The objective is to estimate these coefficients from measurements of $x(t)$ usually at discrete points. The representation is different from a Fourier type form since both amplitudes and exponents are to be calculated (Fourier methods calculate amplitudes for a fixed frequency). The advantage of the Prony series is that short data lengths may be analysed even for 'almost periodic' phenomena. The disadvantage is that the computational procedure involved is nonlinear, involving the solution of a polynomial equation.

The original idea dates back to 1795 but interest was re-awakened in the 1960's. A recent paper by Bento Coelho [26] describes the method and applies it to problems in duct acoustics with comparisons with Fourier methods. The effects of additive noise on the effectiveness of the algorithm is also discussed.

A whole host of parameter estimation schemes have arisen with much of the literature to be found in the control field related to system identification. The article by Lawrence [28] is an overview of identification and parameter estimation. Studies concerned with estimation of structural system parameters (using ARMA models) have been reported by, for example, Gersh et al [29, 30, 22], Prado [31] and Briens [32].

REFERENCES

1. J.P. Burg. Proceedings of the 37th meeting of the Society of Exploration Geophysicists. "Maximum Entropy Spectral Analysis". 1967.
2. J.P. Burg. A new analysis technique for time series data. NATO Advanced Study Institute on Signal Processing with emphasis on Underwater Acoustics. August 1968.
3. D.G. Childers (Editor). Modern spectral analysis. IEE Press Selected Reprint Series. Wiley and Sons Inc. 1978.
4. P.R. Gutowski, E.A. Robinson and S. Treitel. Spectral estimation: fact or fiction. IEEE Trans. Geosci. Electron. Vol. GE-16, pp. 80-84. April 1978.
5. M.B. Priestley. Nonlinear models in time series analysis. Proc. of the Int. Meeting on Time Series Analysis, University of Cambridge, July 1978. Also Course Notes 'Applications of Time Series Analysis', ISVR, Southampton University, 1980.
6. L. Ljung and K. Glover. Frequency domain versus time domain methods in system identification - a brief discussion. Proc. of the 5th IFAC Symposium on Identification and System Parameter Estimation Darmstadt. Ed. R. Isermann. Pergamon Press, 1979.
7. M. Kaveh and G.R. Cooper. An empirical investigation of the properties of the autoregressive spectral estimator. IEEE Trans. Inform Theory, Vol. IT-22, pp. 313-323. May 1976.
8. T.J. Ulrych and T.N. Bishop. Maximum entropy spectral analysis and autoregressive decomposition. Rev. Geophysics and Space Phys. Vol. 13, pp. 183-200. February 1975.
9. J.G. Ables. Maximum entropy spectral analysis. Astron.Astrophys. Suppl. Series. Vol. 15, pp. 383-393. 1974.
10. A. Van den Bos. Alternative interpretation of maximum entropy spectral analysis. IEEE Trans. on Information Theory. July 1971.
11. H. Tong. More on autoregressive model fitting with noisy data by Akaike's information criterion. IEEE Trans. Information Theory, Vol. IT-23, pp. 409-410. May 1977.
12. J.F. Claerbout. Fundamentals of Geophysical Data Processing. Mcgraw Hill Inc. 1976.
13. J. Makhoul. Linear prediction: a tutorial review. Proc. IEEE Vol. 63, pp. 561-580. April 1975.
14. D.J. Laight. Maximum entropy analysis: a comparative study. M.Sc. Thesis, ISVR, 1981.
15. N. Beamish and M.B. Priestley. A study of AR and window spectral estimation. Tech. Report 104, UMIST, Dept. of Maths (Statistics). 1979.
16. H. Akaike. Spectrum estimation through parametric model fitting. IUTAM Symposium Stochastic Problems in Dynamics. Ed. B.L. Clarkson. Pitmans 1976.
17. A. Baggeroer. Confidence intervals for regression (MEM) spectral estimates. IEEE Trans. Inform. Theory, Vol. IT-22, pp. 534-545. September 1976.
18. L. Marple. Resolution of conventional, Fourier, AR and special ARMA methods of spectral analysis. Proc. IEEE Int. Conf. ASSP, Hartford, Conn. 1977.

19. T. Dyson and S. Rao. Some detection and resolution properties of maximum entropy spectral analysis. Signal Processing 2, pp. 261-270. 1980.

20. T.S. Durrani and R.G. Taylor. High resolution spectral estimation. Proc. IEEE Int. Conf. ASSP, Hartford, Conn. 1977.

21. J.K. Vandiver and R.B. Campbell. Estimation of natural frequencies and damping ratios of three similar off-shore platforms using maximum entropy spectral analysis. ASCE Spring Convention, Boston, Mass. 1979.

22. W. Gersch and D.R. Sharpe. Estimation of power spectra with finite order AR models. IEEE Trans. Autom. Control, AC-18, pp. 367-369. 1973.

23. E. Parzen. Some recent advances in time series modelling. IEEE Trans. Autom. Control, Vol. AC-19, pp. 723-730. December 1974.

24. R.H. Jones. Autoregression order selection. Geophysics, vol. 41, pp. 771-773. August 1976.

25. H. Akaike. A new look at the statistical model identification. IEEE Trans. Autom. Control. Vol. AC-19, pp. 716-723. December 1974.

26. J.L. Bento Coelho. Prony and Fourier methods applied to duct acoustics. Proc. Inst. of Ac. Autumn Conference, 1980.

27. P.E. Wellstead. Non parametric methods of system identification. 5th IFAC Symposium on Identification and Parameter Estimation. Ed. R. Isermann, Darmstadt. Pergamon Press, Sept. 1979.

28. P. Lawrence. Modelling, identification and parameter estimation - an overview. Time Series Course Notes, ISVR, Southampton.

29. W. Gersch, N. Nielsen and H. Akaike. Maximum likelihood estimation of structural parameters from random vibration data. Journal of Sound and Vibration, 31(3), 295-308. 1973.

30. W. Gersch and D.A. Foutch. Least squares estimates of structural system parameters using covariance function data. IEEE Trans. Auto. Cont. Vol. Ac-19, No. 6, 1974.

31. G. Prado. Identification of structural parameters using input-output models. Proc. ASME Winter Annual Meeting, Chicago. 1980.

32. P.H. Briens. Time domain recursive estimation of the parameter of vibrating systems. M.Sc. Thesis. ISVR, Southampton University. 1980.

33. R.N. McDonough. Maximum-entropy spatial processing of array data. Geophys. Vol. 39. Dec. 1974.

Vibration transmission and sound radiation

M. Heckl
Technische Universität Berlin
Institut für Technische Akustik
Einsteinufer 27, D-1000 Berlin 10

Summary

Waves in solids and in gases or fluids are governed by very similar equations, which are based on continuity of mass, continuity of momentum and an equation of state. Therefore sound radiation and vibration transmission into solid media have much in common. In this lecture radiation from planar sources, reflection at plane boundaries, the behaviour of wave guides, source impedances etc. are compared and discussed.

1. Introduction

In this paper - according to its title - some information is to be given on the relation between vibration transmission and sound radiation. Such a topic can be understood in several different ways; e.g. one might relate the vibrations of a surface to the radiated sound or one might think of the excitation of structural vibrations by sound waves. But these two subjects are thoroughly treated in the lectures by M.C. Junger and B.L. Clarkson; therefore it seemed appropriate in this paper to relate waves and vibrations in solid structures to sound waves in fluids or gases and to show some of the similarities and differences between the two.

The basic reason for the similarities between waves in fluids or gases and waves in solids is well known, it is due to the fact that compressional waves are possible in all of them. The basic reason for the differences is also clear, it is due to the fact that shear waves can exist in solid materials but not in gases or fluids. This latter statement has to be qualified somewhat: it is absolutely correct only if viscosity is neglected; if this is not the case there are some highly damped shear waves near the boundary of two media (especially between a gas or fluid and a solid material). Although these shear waves are extremely important when the sound absorption in porous materials is investigated, they can be neglected here, because they are restricted to the very thin so-called "acoustic boundary layer".

2. Wave equations

All wave equations can be drived from continuity of mass and momentum and an equation describing the properties of the material. If we deal with a material consisting of one component only (i.e. no mixtures and no chemical reactions) continuity of mass and momentum can be expressed by the familiar relations

$$\frac{\partial \rho_t}{\partial t} + \frac{\partial \rho_t v_i}{\partial x_i} = 0 \qquad (1)$$

$$\frac{\partial \rho_t v_i}{\partial t} + \frac{\partial \sigma_{ij}}{\partial x_j} = \frac{\partial \rho_t v_i v_j}{\partial x_j} \,. \qquad (2) \quad i, j = 1, 2, 3$$

Here ρ_t is the density, v_i are the components of the velocity vector, t is the time, x_i are the space components and σ_{ij} is the stress tensor that includes all the forces acting on a volume element. Volume forces from the outside (e.g. variation in gravity) easily could have been included in eq.(2) but since they do not play an important role in practical cases they are omitted here. As usual the summation convention has been applied. If eq.(1) is differentiated with respect to t and if the three equations (2) are differentiated with respect to x_i and added they can be combined to give

$$\frac{\partial^2 \rho_t}{\partial t^2} - \frac{\partial^2 \sigma_{ij}}{\partial x_i \partial x_j} = \frac{\partial^2 \rho_t v_i v_j}{\partial x_i \partial x_j} \,. \qquad (3)$$

Eq.(1)-(3) hold for gases as well as fluids or solids. The differences appear when the stress tensor is related to the components of the displacements or velocities; i.e. when

the equation of state of the material is introduced. Three cases are of interest for this paper /1/:

a. Fluids or gases with vanishing viscosity

$$\sigma_{ij} = p_t \delta_{ij} \tag{4}$$

Here p_t is the scalar pressure and δ_{ij} is the Kronecker symbol.

b. Fluids or gases with viscosity

$$\sigma_{ij} = p_t \delta_{ij} + \mu\left[-\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} + \frac{2}{3}\left(\frac{\partial v_k}{\partial x_k}\right)\delta_{ij}\right] \tag{5}$$

Here μ is the coefficient of viscosity.

c. Solids

$$\sigma_{ij} = p_t \delta_{ij} + G\left[-\frac{\partial \xi_i}{\partial x_j} - \frac{\partial \xi_j}{\partial x_i} + \frac{2}{3}\left(\frac{\partial \xi_k}{\partial x_k}\right)\delta_{ij}\right] . \tag{6}$$

Here G is the shear modulus of the material and ξ_i are the components of the displacement vector. Obviously the velocity v_i is the time derivation of the displacement ξ_i.

Eq.(4)-(6) have in common, that they are linear and that the stresses at a certain time depend on the displacements or velocities at exactly the same time (i.e. there is no influence of past as would be the case in materials with "memory"). With the same order of accuracy we can assume, that changes in pressure are proportional to changes in density. Thus with

$$p_t = \bar{p} + p \quad ; \quad \rho_t = \bar{\rho} + \rho \tag{7}$$

we can write

$$p = K\rho/\bar{\rho} . \tag{8}$$

Here the bar denotes the mean value; i.e. $\partial\bar{p}/\partial x_i = 0$; $\partial\bar{\rho}/\partial x_i = 0$. p and ρ are the deviations of pressure and density from the mean values and K is the modulus of compressibility. If eq.(4)-(8) are introduced into (3) we obtain the following equations for the three cases:

a) $$\frac{\partial^2 \rho_t}{\partial t^2} - \Delta p_t = \frac{\partial^2 \rho}{\partial t^2} - \frac{K}{\bar{\rho}}\Delta\rho = \frac{\partial^2 \rho_t v_i v_j}{\partial x_i \partial x_j} \tag{9}$$

b) $$\frac{\partial^2 \rho}{\partial t^2} - \Delta p - \mu\frac{\partial^2}{\partial x_i \partial x_j}\left[-\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} + \frac{2}{3}\left(\frac{\partial v_k}{\partial x_k}\right)\delta_{ij}\right] = \frac{\partial^2 \rho_t v_i v_j}{\partial x_i \partial x_j} \tag{10}$$

c) $$\frac{\partial^2 \rho}{\partial t^2} - \Delta p - G\frac{\partial^2}{\partial x_i \partial x_j}\left[-\frac{\partial \xi_i}{\partial x_j} - \frac{\partial \xi_j}{\partial x_i} + \frac{2}{3}\left(\frac{\partial \xi_k}{\partial x_k}\right)\delta_{ij}\right] = \frac{\partial^2 \rho_t v_i v_j}{\partial x_i \partial x_j} . \tag{11}$$

(Here Δ is the Laplacian operator)

The first of these expressions (eq.9), which may be called Lighthill's equation /2/ (or Lighthill analogy), is well known because it is used in almost all theoretical flow noise problems. In the purely linear case, i.e. when $\rho v_i v_j < p$ (that is when the flow velocities are very small compared to the speed of sound) the right hand side of eq.(9) can be neglected and the normal sound wave equation is obtained. The wave speed is in this case

$$c_s = \sqrt{K/\bar{\rho}} . \tag{12}$$

The second equation (eq.10) - with the nonlinear term on the right hand side omitted - is used to describe the sound field in viscous media and the details of the wave motion near a solid surface inside the acoustic boundary layer.

In order to bring eq.(11) into a familiar form the nonlinear term is again neglected and the other terms are rearranged. This way it is found that for an elastic isotropic solid material the relation

$$\frac{\partial^2 \rho}{\partial t^2} - \Delta p + \frac{4}{3}G\Delta\left(\frac{\partial \xi_i}{\partial x_i}\right) = 0 \tag{13}$$

holds

Introducing the divergence

$$\delta = \frac{\partial \xi_i}{\partial x_i} = \frac{\partial \xi_1}{\partial x_1} + \frac{\partial \xi_2}{\partial x_2} + \frac{\partial \xi_3}{\partial x_3} \tag{14}$$

and using the fact that δ is just the relative change of a volume element we can write

$$\rho/\bar{\rho} = -\delta \tag{15}$$

and using eq.(8)

$$p = -K\delta . \tag{16}$$

This way we get

$$\bar{\rho}\frac{\partial^2 \delta}{\partial t^2} - \left(K + \frac{4}{3}G\right)\Delta\delta = 0 . \tag{17}$$

If the modulus of compressibility and the shear modulus are expressed in terms of Young's modulus E and Poisson's number ν; i.e. if we write /3/

$$K = \frac{E}{3(1-2\nu)} \quad ; \quad G = \frac{E}{2(1+\nu)} \tag{18}$$

eq.(17) becomes

$$\frac{\partial^2 \delta}{\partial t^2} = \frac{E}{\bar{\rho}} \frac{1-\nu}{(1+\nu)(1-2\nu)} \Delta\delta = c_c^2 \Delta\delta . \tag{19}$$

Here

$$c_c = \sqrt{\frac{E}{\bar{\rho}} \frac{1-\nu}{(1+\nu)(1-2\nu)}} \tag{20}$$

is the speed of propagation for compressional waves (in this paper the term longitudinal waves is purposely avoided because there is always the danger that there may be a confusion with the longitudinal waves in bars which have a speed $c_L = \sqrt{E/\bar{\rho}}$ and which show some lateral contraction).

There are two types of solution for eq.(19). The first type consists of non-vanishing functions of $\delta(x_i,t)$ which fulfill eq.(19) (and the prescribed boundary conditions); the other is the "trivial" solution $\delta = 0$. In order to find out wether it has any physical significance, we introduce it into the basic eq.(1), (2), (6).
Since for $\delta = 0$ we have

$$\frac{\partial \xi_{Ti}}{\partial x_i} = 0 \; ; \quad \rho_t = \bar{\rho} \; ; \; \frac{\partial \rho_t}{\partial t} = 0 ; \quad p = 0$$

we find

$$\bar{\rho}\frac{\partial v_{Ti}}{\partial t} + \frac{\partial \sigma_{ij}}{\partial x_j} = \bar{\rho}\frac{\partial v_{Ti} v_{Tj}}{\partial x_j} \tag{21}$$

$$\sigma_{ij} = -G\left(\frac{\partial \xi_{Ti}}{\partial x_j} + \frac{\partial \xi_{Tj}}{\partial x_i}\right) . \tag{22}$$

Within the range of linear approximations, that will be used throughout the rest of the paper we can neglect the $v_i v_j$ term and we can set $v_i = \partial \xi_i / \partial t$. Thus we get

$$\bar{\rho}\frac{\partial^2 \xi_{Ti}}{\partial t^2} = G\frac{\partial}{\partial x_j}\left(\frac{\partial \xi_{Ti}}{\partial x_j} + \frac{\partial \xi_{Tj}}{\partial x_i}\right) \tag{23}$$

(The index T indicates that this is only that part of the full solution for which $\delta = 0$ i.e. which has vanishing divergence). Eq.(23) can be cast into a familiar form if it is written with index K instead of i; thus

$$\bar{\rho}\frac{\partial^2 \xi_{TK}}{\partial t^2} = G\frac{\partial}{\partial x_j}\left(\frac{\partial \xi_{TK}}{\partial x_j} + \frac{\partial \xi_{Tj}}{\partial x_K}\right) . \tag{24}$$

Differentiating eq.(23) with respect to x_K and eq.(24) with respect to x_i and subtracting gives

$$\bar{\rho}\frac{\partial^2}{\partial t^2}\left(\frac{\partial \xi_{Ti}}{\partial x_K} - \frac{\partial \xi_{TK}}{\partial x_i}\right) = G\Delta\left(\frac{\partial \xi_{Ti}}{\partial x_K} - \frac{\partial \xi_{TK}}{\partial x_i}\right) . \tag{25}$$

Eq.(25) is correct for all values of i and κ but it makes sense only for $i \neq \kappa$. Of these combinations only the following three are of interest

$$\psi_1 = \frac{\partial \xi_{T3}}{\partial x_2} - \frac{\partial \xi_{T2}}{\partial x_3} \; ; \quad \psi_2 = \frac{\partial \xi_{T1}}{\partial x_3} - \frac{\partial \xi_{T3}}{\partial x_1} \qquad (26)$$

$$\psi_3 = \frac{\partial \xi_{T2}}{\partial x_1} - \frac{\partial \xi_{T1}}{\partial x_2} \; .$$

Thus we finally have

$$\frac{\partial^2 \psi_i}{\partial t^2} = \frac{G}{\bar{\rho}} \Delta \psi_i = c_T^2 \Delta \psi_i \; . \qquad (27)$$

Here $c_T = \sqrt{G/\bar{\rho}}$ is the speed of propagation for shear waves.

Since δ is just the divergence of the displacement vector ξ_i and Ψ_i the components of its curl, the two eq.(19) and (27) give the complete solution. Therefore the usual way of solving problems in the field of vibration transmission is to write

$$\vec{\xi} = \operatorname{grad} A + \operatorname{curl} \vec{B} \qquad (28)$$

giving $\varphi = \Delta A$ and $\psi = \operatorname{curl} \operatorname{curl} \vec{B}$ then the wave equations - taking into account the boundary conditions which also have to be expressed in terms of A and B - are solved using standard techniques and the two parts are combined to give the whole displacement field.

In this paper, however, an alternate approach will be used which in some books - to this author's surprise - is called more cumbersome. In this approach the displacement is obtained directly without the detour using the divergence and the curl. The starting equation for the method is obtained by combining eq.(1) and (2) to give

$$\frac{\partial \rho_t v_i}{\partial t} + \frac{\partial \sigma_{ij}}{\partial x_j} = \rho_t \frac{\partial v_i}{\partial t} + v_i \frac{\partial \rho_t}{\partial t} + \frac{\partial \sigma_{ij}}{\partial x_j} = \rho_t \frac{\partial v_i}{\partial t} - v_i \frac{\partial \rho_t v_j}{\partial x_j} + \frac{\partial \sigma_{ij}}{\partial x_j} = - \frac{\partial \rho_t v_i v_j}{\partial x_j}$$

or

$$\rho_t \frac{\partial v_i}{\partial t} + \frac{\partial \sigma_{ij}}{\partial x_j} = v_i \frac{\partial \rho_t v_j}{\partial x_j} - \frac{\partial \rho_t v_i v_j}{\partial x_j} \; . \qquad (29)$$

Within the framework of linear approximations we can put $\rho_t = \bar{\rho}$ and neglect the second order terms on the right hand side.

Thus together with eq.(6) and (16) we obtain

$$\bar{\rho} \frac{\partial v_i}{\partial t} = \bar{\rho} \frac{\partial^2 \xi_i}{\partial t^2} = - \frac{\partial}{\partial x_j} \left[\left(-K + \frac{2}{3} G \right) \frac{\partial \xi_\kappa}{\partial x_\kappa} \delta_{ij} - G \left(\frac{\partial \xi_i}{\partial x_j} + \frac{\partial \xi_j}{\partial x_i} \right) \right] . \qquad (30)$$

After some rearrangements this becomes

$$\bar{\rho} \frac{\partial^2 \xi_i}{\partial t^2} = G \Delta \xi_i + \left(K + \frac{1}{3} G \right) \frac{\partial}{\partial x_i} \left(\frac{\partial \xi_\kappa}{\partial x_\kappa} \right) . \qquad (31)$$

If here K is replaced by G or ν (see eq.(18)) the result is

$$\frac{1}{c_T^2} \frac{\partial^2 \xi_i}{\partial t^2} = \Delta \xi_i + \frac{1}{1-2\nu} \frac{\partial}{\partial x_i} \left(\frac{\partial \xi_\kappa}{\partial x_\kappa} \right) . \qquad (32)$$

Obviously this equation also must break down into a shear wave part and into a compressional wave part. This can be shown rather easily if everything is expressed in terms of Fourier transforms. Thus if we write

$$\xi_i(x_i, t) = \int_{-\infty}^{+\infty} \check{\xi}_i(k_i, \omega) e^{-jk_i x_i} e^{j\omega t} \, dk_1 dk_2 dk_3 d\omega \qquad (33)$$

Eq.(32) reduces to

$$-k_T^2 \check{\xi}_i = -k^2 \check{\xi}_i - \frac{1}{1-2\nu} k_i \left(\check{\xi}_\kappa k_\kappa \right) \qquad (34)$$

Here we introduced the shear wave number

$$k_T^2 = \omega^2/c_T^2 \tag{35}$$

and the abreviation

$$k^2 = k_1^2 + k_2^2 + k_3^2 . \tag{36}$$

Eq.(34) is a system of three linear equations, it can be solved only if the determinant becomes zero. This leads to a polynomial that vanishes when either

$$k^2 = k_T^2 \tag{37}$$

or

$$k^2 = \frac{1-2\nu}{2-2\nu} k_T^2 = k_L^2 . \tag{38}$$

These are just the wave numbers for shear waves and compressional waves.

Thus it follows from the linear system (34) that the general solution is

$$\xi_i(x_i,t) = \int_{-\infty}^{+\infty}\left[\check{\xi}_{iL}(k_1,k_2,\omega)e^{\pm jq_L x_3} + \check{\xi}_{iT}(k_1,k_2,\omega)e^{\pm jq_T x_3}\right]e^{-jk_1x_1}e^{-jk_2x_2}e^{j\omega t}\,dk\,d\omega \tag{39}$$

Here

$$q_T = \sqrt{k_T^2 - k_1^2 - k_2^2} \tag{40}$$

is the value of k_3 that satisfies eq.(37) and

$$q_L = \sqrt{k_L^2 - k_1^2 - k_2^2} \tag{41}$$

is the value of k_3 that satisfies eq.(38). Furthermore eq.(34) yields the following relations

$$\check{\xi}_{3L} = \frac{q_L}{k_1}\check{\xi}_{1L} \; ; \quad \check{\xi}_{2L} = \frac{k_2}{k_1}\check{\xi}_{1L} \; ; \quad k_1\check{\xi}_{1T} + k_2\check{\xi}_{2T} + q_T\check{\xi}_{3T} = 0 \tag{42}$$

This might also be written as

$$\check{\xi}_{iL}k_j = \check{\xi}_{jL}k_i \; ; \quad \check{\xi}_{iT}k_i = 0. \tag{42a}$$

This way one is just left with enough free variables to fulfill the boundary conditions.

3. Sound radiation and vibration transmission from planar sources

3.1 Gases or fluids

If the sound wave equation (9) is written without the nonlinear term and if the density change is expressed through the fluctuating pressure (eq.8) it reads

$$\Delta p + k_s^2 p = 0. \tag{43}$$

Here we have restricted ourselves to purely harmonic motion with angular frequency ω (the factor $e^{j\omega t}$ is omitted throughout) and we have introduced the wave number for sound waves

$$k_s = \omega/c_s . \tag{44}$$

In addition we need the linearised version of the equation of conservation of momentum. It follows from eq.(2) and (4) as

$$\bar{\rho}\frac{\partial v_i}{\partial t} = -\frac{\partial p}{\partial x_i} \qquad j\omega\bar{\rho}v_i = -\frac{\partial p}{\partial x_i} . \tag{45}$$

If all field quantities are expressed in terms of spatial Fourier transforms (similar to Eq.(33)) the only information we get out of the wave equation (43) is that

$$k_1^2 + k_2^2 + k_3^2 = k_s^2 . \tag{46}$$

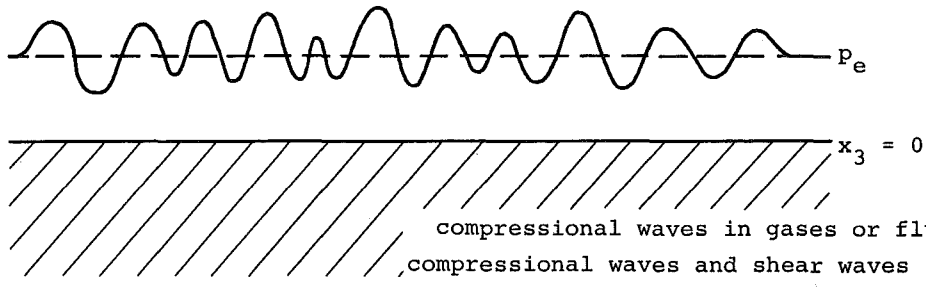
Thus if we rename k_3 as q_s we have

$$q_s = \sqrt{k_s^2 - k_1^2 - k_2^2} . \tag{47}$$

The sound pressure field is therefore given by

$$p(x_i) = \int_{-\infty}^{+\infty}\left[\check{p}_A(k_1,k_2)e^{jq_s x_3} + \check{p}_R(k_1,k_2)e^{-jq_s x_3}\right]e^{-jk_1x_1}e^{-jk_2x_2}\,dk_1\,dk_2 . \tag{48}$$

If we take as a first example the sound field in a semi-infinite space that is excited at its surface by a given pressure distribution $p_e(x_1,x_2)$ which has to be specified over the entire surface - e.g. an air-borne sound wave hitting a free water surface - we can use the usual radiation condition which states that no energy comes from infinity. This gives in our case $\check{p}_R(k_1,k_2) = 0$. The remaining unknown function $\check{p}_A(k_1,k_2)$ is determined



by the fact that at the boundary $x_3 = 0$ there is continuity of pressure; i.e.

$$p_e(x_1,x_2) = \int \check{p}_A(k_1,k_2)\, e^{-jk_1x_1} e^{-jk_2x_2}\, dk_1\, dk_2. \qquad (49)$$

Using the orthogonality property of the exponential function eq.(49) can be inverted to give

$$\check{p}_A(k_1,k_2) = \frac{1}{4\pi^2} \int p_e(x_1,x_2)\, e^{jk_1x_1} e^{jk_2x_2}\, dx_1\, dx_2. \qquad (50)$$

Now the problem is solved provided we chose the sign of q_S in such a way that no waves come from infinity. For the example shown above where x_3 is always negative this would mean that q_S has to be positive when it is real and negative when it is imaginary.

In Fig.(1) the paths of the fluid particles that have been calculated using the above equations can be seen for some examples. The results shown in the figure are based on two dimensional calculations (i.e. variable x_2 does not appear).

By looking at the figure (and many others, including a film) the following general conclusions can be drawn:

a. If the excitation has a spatial periodicity with a wave length which is smaller than the wave length of the sound waves in the medium in the region $x_3<0$ there exists a strong near field. The radiated sound energy is small, it seems to come from those points where the periodicity is not perfect; especially from the ends of the excited area. In this case the radiated sound depends rather strongly on the details of the excitation near the ends.

b. The radiation is rather strong when the wave length of the excitation coincides with the sound wave length. The amplitude in this case depends on the number of wave lengths that fit into the excited area.

c. If the excitation is due to a convected pattern, no sound energy is generated (i.e. there are no propagating waves) provided the pattern moves with a speed which is below the speed of sound. This statement holds for any pattern irrespective of its shape provided the shape remains constant.

d. If the excited area is small compared to the wavelength of sound, there is a strong nearfield. If the pressure within this region is constant the farfield depends on the product of pressure times area; the nearfield is determined by the details of the source configuration.

e. The strength of the sound wave which travels perpendicular to the free surface is determined by the net pressure averaged over the whole surface.

Similar calculations can be made and similar pictures could have been obtained for the sound radiation if the velocity over the surface $x_3 = 0$ were prescribed; e.g. when the sound is generated by a planar radiator in a baffle (see lectures by M.C. Junger). All that has to be done is to express the velocity in terms of the pressure by using eq.(45) and to apply the boundary conditions. If for example the velocity $v_e(x_1,x_2)$ is given in the plane $x_3 = 0$ the Fourier transform of the sound pressure is $\check{p}_A(k_1,k_2) = \check{v}_e(k_1,k_2) \cdot \omega/q_S$. When this expression is inserted into eq.(48) (with $\check{p}_R = 0$) the sound field is obtained. In spite of the fact that q_S can become zero the solution is well behaved except when the exciting pattern is purely periodic with wavenumber $k_1^2 + k_2^2 = k_S^2$. A more complicated case is the radiation when on parts of the area the pressure and on other parts the velocity is prescribed. For such (mixed boundary value) problems more elaborate mathematical techniques (Wiener-Hopf) have to be used.

3.2 Solids

For solids the same basic procedure can be used as for gases or liquids. The basic equation in this case is eq.(32). Expressing all quantities in terms of their Fourier transforms and restricting ourselves to purely harmonic motion gives the following general solution according to eq.(39)

$$\xi_i(x_i) = \int\int \left[\check{\xi}_{iAC} e^{j q_C x_3} + \check{\xi}_{iRC} e^{-j q_C x_3} + \check{\xi}_{iAT} e^{j q_T x_3} + \check{\xi}_{iRT} e^{-j q_T x_3} \right] e^{-j k_1 x_1} e^{-j k_2 x_2} \, dk_1 \, dk_2 . \tag{51}$$

The quantities q_C and q_T are given by eq.(40) and (41) and the relations (42) have to be observed.

In addition to eq.(51) we need expressions relating the stresses and the displacements. Obviously these are eq.(6) and (16) which in terms of Fourier transforms become

$$\frac{\check{\sigma}_{ij}}{jG} = \frac{2\nu}{1-2\nu} \delta_{ij} (k_K \check{\xi}_{KC}) + 2 k_j \check{\xi}_{jC} + k_j \check{\xi}_{iT} + k_i \check{\xi}_{jT} . \tag{52}$$

Here the relations (42a) have been employed already, but they are needed again to eliminate some of the variables ξ. Eq.(52) is a system of nine equations but it is highly redundant; e.g. $\check{\sigma}_{ij} = \check{\sigma}_{ji}$. Out of the remaining six expressions only three are needed. For the geometry shown in sect. 3.1 these are, after $\check{\xi}_{1C}$, $\check{\xi}_{2C}$ and $\check{\xi}_{1T}$ have been eliminated

$$\frac{\check{\sigma}_{33}}{jG} = \frac{1}{q_C} \left[k_T^2 - 2(k_1^2 + k_2^2) \right] \check{\xi}_{3C} + 2 q_T \check{\xi}_{3T} \tag{53}$$

$$\frac{\check{\sigma}_{23}}{jG} = 2 k_2 \check{\xi}_{3C} + k_2 \check{\xi}_{3T} + q_T \check{\xi}_{2T} \tag{54}$$

$$\frac{\check{\sigma}_{13}}{jG} = 2 k_1 \check{\xi}_{3C} + (k_1^2 - q_T^2) \frac{\check{\xi}_{3T}}{k_1} - \frac{k_2 q_T}{k_1} \check{\xi}_{2T} . \tag{55}$$

By combining these three equations one finally gets

$$\check{\xi}_{3C} = \frac{-1}{jG} \frac{\check{\sigma}_{33} \left[2(k_1^2 + k_2^2) - k_T^2 \right] - \check{\sigma}_{32} 2 k_2 q_T - \check{\sigma}_{31} 2 k_1 q_T}{\left[2(k_1^2 + k_2^2) - k_T^2 \right]^2 + 4 q_C q_T (k_1^2 + k_2^2)} q_C \tag{56}$$

and

$$\check{\xi}_{3T} = \frac{1}{jG} \frac{\check{\sigma}_{33} 2 q_C (k_1^2 + k_2^2) + (\check{\sigma}_{32} k_2 + \check{\sigma}_{31} k_1) \left[2(k_1^2 + k_2^2) - k_T^2 \right]}{\left[2(k_1^2 + k_2^2) - k_T^2 \right]^2 + 4 q_T q_C (k_1^2 + k_2^2)} \tag{57}$$

Thus if the Fourier transforms of the normal force σ_{33} and the shear forces σ_{32} and σ_{31} in a plane are known the Fourier transforms of the displacements $\check{\xi}_{3C}$ and $\check{\xi}_{3T}$ can obtained from eq.(56) and (57). Since the other displacement components follow from eq.(42) or (42a) the problem is solved. If instead of the forces the displacements are given the procedure is similar. The problem can also easily be solved if in one entire x_3-plane one component of the displacement vector and two force components are prescirbed. If for example $\xi_e = \xi_{3C} + \xi_{3T}$ and σ_{23} and σ_{13} are known one finds

$$\check{\xi}_{3C} = -\frac{2(k_1^2 + k_2^2) - k_T^2}{k_T^2} \check{\xi}_e + \frac{1}{jG} \frac{k_2 \check{\sigma}_{23} + k_1 \check{\sigma}_{13}}{k_T^2} \; ; \; \check{\xi}_{3T} = 2 \frac{k_1^2 + k_2^2}{k_T^2} \check{\xi}_e - \frac{1}{jG} \frac{k_2 \check{\sigma}_{23} + k_1 \check{\sigma}_{13}}{k_T^2} \tag{58}$$

The other components can again be found from eq.(42a).

In Fig.2 the paths of the particles that have been calculated using the above equations can be seen for examples, which again are of two-dimensional nature.

From these figures and many other similar data some general conclusions can be drawn which are very similar to those at the end of section 3.1.

a. If the excitation has a spatial periodicity with wavelength λ_e, the generation of compressional waves is weak if $\lambda_e < \lambda_C$ (λ_C being the free compressional wavelength $\lambda_C = 2\pi/k_C$); in addition the generation of shear waves is weak if $\lambda_e < \lambda_T = 2\pi/k_T$. Deviations from the periodicity in the excitation have the same effect as described

in sect. 3.1.

b. The generation of compressional waves is very strong when $\lambda_e \approx \lambda_C$. The generation of shear waves is very strong when $\lambda_e \approx \lambda_T$. The amplitudes depend mainly on the size of the excitated area.

c. If the excitation is due to a "frozen" pattern that moves with convection speed U_e no compressional waves are propagated if $U_e < C_C$ and no shear waves are propagated if $U_e < C_T$.

d. If the excited area is very small conclusion d of sect. 3.1 holds, too.

e. Conclusion e of sect. 3.1 holds also.

One peculiarity of wave propagation in solids with a free surface is the existence of Rayleigh waves. These are waves that are limited to a layer near the free surface (see Fig.2). The particle path in this case is neither parallel nor perpendicular to the direction of wave propagation; instead the particles move along ellipses. The wave number of Rayleigh waves is determined by the vanishing of the denominator of eq.(56) or (57). This leads to a third order polynomial in $(k_1^2+k_2^2)$ which has two complex and one real root. It can be shown that this real root which we call k_R^2 is slightly bigger than k_T^2. It is approximately given by

$$k_R^2 \approx k_T^2 \left(1 + \frac{\sqrt{\frac{1}{2}-\nu+\frac{3}{2}\nu^2} - \nu}{2(1-\nu)}\right). \qquad (59)$$

As in the case of the other wave types the generation of Rayleigh waves is especially pronounced when $\lambda_e > \lambda_R = 2\pi/k_R$ or when $U_e > C_R = \omega/k_R$.

4. Reflection of waves

4.1 Gases or fluids

When sound waves are reflected at a plane boundary the resulting field can be understood as the direct field combined with the field of an image source. The amplitude and the phase of the image source depend on the reflection coefficient of the boundary. Since this phenomenon is so simple no examples need to be given here.

4.2 Solids

For waves in solids things are somewhat more complicated and usually it is not possible to express the resulting sound field in terms of image sources. The reason for this is that there is wave type conversion at boundaries; i.e. compressional waves are partially transfered to shear waves and vice versa.

Using the Fourier transform method again it is more or less an exercise in linear algebra to calculate the reflection when the incoming wave and the properties of the boundaries are known.

Out of the many possible combinations of displacements and stresses, we shall treat here only the case when the displacements of the incoming waves are known and the stresses σ_{33}, σ_{23}, σ_{13} at the surface $x_3 = 0$ are given (usually they are zero). For this case the incoming wave may be written

$$\xi_{iA}(x_i) = \int\left[\check{\xi}_{iAC}\, e^{-jq_C x_3} + \check{\xi}_{iAT}\, e^{-jq_T x_3}\right] e^{-jk_1 x_1}\, e^{-jk_2 x_2}\, dk_1\, dk_2. \qquad (60)$$

Not all of the functions ξ_{iAC} and ξ_{iAT} need be given. It is enough to know two of them the rest follows from the relations (42a).

The reflected wave has the same wave numbers in x_1 and x_2 direction (trace-matching /4/) but in x_3-direction the wave numbers are opposite. Thus we have for the reflected wave:

$$\xi_{iR}(x_i) = \int\left[\check{\xi}_{iRC}\, e^{jq_C x_3} + \check{\xi}_{iRT}\, e^{jq_T x_3}\right] e^{-jk_1 x_1}\, e^{-jk_2 x_2}\, dk_1\, dk_2, \qquad (61)$$

If eq.(60) and (61) are introduced into eq.(53),(54),(55) the following expressions are obtained

$$\check{\xi}_{3RC} = \frac{\frac{1}{jG}\left[N q_c \check{\sigma}_{33} + 2 q_T q_c (k_2 \check{\sigma}_{23} + k_1 \check{\sigma}_{13})\right] + \left[N^2 - 4 q_T q_c (k_1^2 + k_2^2)\right] \check{\xi}_{3AC} - 4N q_T q_c \check{\xi}_{3AT}}{N^2 + 4 q_c q_T (k_1^2 + k_2^2)} \tag{62}$$

$$\check{\xi}_{3RT} = \frac{\frac{1}{jG}\left[-2 q_c (k_1^2 + k_2^2) \check{\sigma}_{33} + N(k_2 \check{\sigma}_{23} + k_1 \check{\sigma}_{13})\right] - 4N(k_1^2 + k_2^2) \check{\xi}_{3AC} + \left[4 q_T q_c (k_1^2 + k_2^2) - N^2\right] \check{\xi}_{3AT}}{N^2 + 4 q_c q_T (k_1^2 + k_2^2)} \tag{63}$$

Here the abbreviation

$$N = 2(k_1^2 + k_2^2) - k_T^2 \tag{64}$$

has been used.

Thus if the incoming wave and the three stresses σ_{33}, σ_{23}, σ_{13} are known (most often they are zero because this corresponds to a free surface) the reflected waves can be calculated.

The wave field in solids when one reflection has taken place was also calculated for several examples.Surprisingly it turned out that quite often the wavefields look very similar to those in fluids or gases in spite of the fact that the reflection mechanism is rather different.

5. Wave guides

If we now turn to media with two or more parallel sides, we obtain wave guides. Here again the spatial Fourier transform method is useful provided that everything can conveniently be expressed in cartesion coordinates. (For some other simple geometries similar calculations can be performed by using other coordinate systems). All that has to be done is to express the field in terms of plane waves in opposite directions and to adjust the freee variables in such a way that the boundary conditions are fulfilled. If this is done, it turns out, that there are a number of wave types that are propagated unchanged if the frequency is above a certain cut-off value.

5.1 Gases or fluids

A very simple wave guide consists of a gas-filled duct with rigid walls. In this case the normal component of the velocity vanishes at $x_3 = \pm h$. Since this corresponds to

$$v_3 = -\frac{1}{j\omega}\frac{\partial p}{\partial x_3} = 0, \quad \text{for } x_3 = \pm h \tag{65}$$

eq.(48) leads to $\sin 2q_S h = 0$ or $q_S h = \pi n/2$.

Taking into account that $\mathrm{Re}\{q_S\} \leq k_S$, it turns out that propagating waves with n nodal lines in x_3-direction are possible only if

$$\frac{\omega}{c_S} = k_S \geq n\pi/2h \qquad n = 0,1,2,.... \tag{66}$$

Another simple wave guide consist of a fluid-filled duct with soft or free surfaces. In this case the pressure disappears at $x_3 = \pm h$ and it turns out that propagating waves are possible only if

$$\frac{\omega}{c_S} = k_S \geq \frac{2n+1}{2}\frac{\pi}{h}.$$

Fig.3 shows some examples of wave patterns in fluid- or gas-filled ducts.

5.2 Solids

For solid, rectangular wave guides the starting equations are eq.(51)-(55). The mathematical procedure is the same that has been applied in sect. 5.1 for gases and fluids, the only difference is that the calculations are lengthier.

If we restrict ourselves to free beams and plates with thickness 2 h; i.e. to wave guides with $\sigma_{33} = \sigma_{32} = \sigma_{31} = 0$ at the planes $x_3 = \pm h$, eq.(51) introduced into (53)-(55) yields the following conditions for free wave propagation.

$$N^2 \operatorname{tg} q_c h + 4(k_1^2+k_2^2) q_c q_\tau \operatorname{tg} q_\tau h = 0 \qquad (67)$$

$$N^2/\operatorname{tg} q_c h + 4(k_1^2+k_2^2) q_c q_\tau/\operatorname{tg} q_\tau h = 0 \qquad (68)$$

(See eq.(64) for N).

Eq.(67) and (68) have an infinite number of solutions, but unfortunately they cannot be expressed explicitely. For many practical applications it is however enough to know the lowest order solution. They are found by expanding the tg-function up to the third order. The result is that free waves on thin plates are possible if

$$(k_1^2+k_2^2)^2 = \omega^2 \frac{\bar{\rho}}{E} \frac{3(1-\nu^2)}{h^2} = k_B^4 \qquad (69)$$

$$(k_1^2+k_2^2) = \omega^2 \frac{\bar{\rho}}{E}(1-\nu^2) = k_L^2 . \qquad (70)$$

Eq.(69) which follows from (67) is just the dispersion relation for bending waves on plates, eq.(70) which is derived from (68) describes (quasi-) longitudinal waves on plates.

Fig.4 shows some examples of the vibration patterns of plates of different thickness. Fig.5 shows plates that are excited over a small area with a constant pressure. In this case the stress component σ_{33} is zero except for a small area over which it is constant. By Fourier-transforming σ_{33} and applying eq.(51)-(55) the wave-patterns are obtained.

6. Impedances

For problems of sound radiation and of vibration transmission it is important to know the impedance or admittance of a system. This way the mechanical power generated by a source and the transmission properties at discontinuities or junctions can be treated in a rather convenient way. The definition of impedance Z and admittance A are

$$Z = \frac{\hat{F}}{\hat{v}} \; ; \quad A = \frac{\hat{v}}{\hat{F}} \qquad (71)$$

where $\hat{F}$ is the complex amplitude of the force and $\hat{v}$ the complex amplitude of the velocity; obviously Z and A are defined only when the motion is purely harmonic in time.

Quite often it is said that F is a "point source" i.e. a force acting over a vanishing small area. Such a condition, however, causes some difficulties, because if the force is kept constant and the area is made smaller and smaller the velocity around the source region increases. This is true for all materials with finite shear stiffness (i.e. for all solids) and it is especially true for liquid or gases. Thus it turns out that the concept of a point impedance is not a very useful one if it is taken literally because it depends on the size of a "point". The problem can be resolved, however, if the discussion is restricted to the real part of admittance. This quantity determines the mechanical power which is transmitted from a source into a medium. If damping is not present it must be possible to find this power again in the field. In the farfield, however, the wave numbers k_1 and k_2 have to be below a certain limit otherwise the quantities q_S or q_C or q_T become imaginary which corresponds to exponentially decaying near fields. Thus the integration of eq.(51) has to be performed only in the range $k_1^2+k_2^2<k_S^2$ or k_C^2 or k_T^2. If we now take an excitation of the form

$$\sigma_{33}(x_1,x_2) = \begin{cases} F/\ell^2 & \text{within } x_1^2+x_2^2<\ell^2 \\ 0 & \text{otherwise} \end{cases} \qquad (72)$$

- i.e. the excitation is restricted to a region of dimension l - we get a Fourier-transform $\check{\sigma}_{33}(k_1,k_2)$ which is constant for $k_1^2+k_2^2<1/l^2$. Thus if $k_S l<1$ or $k_T l<1$ the integration in the far field extends only over those values of k_1 and k_2 for which $\check{\sigma}_{33}(k_1,k_3)$ is constant; therefore we get a result which is independent of the details of the source configuration.

So the result of the calculation is that the real part of the admittance is indepen-

dent of the source dimensions and source shape as long as $k_S l<1$ or in the case of solids $k_T l<1$. The imaginary part of the admittance depends very strongly on the source configuration. If the source is very small there is a very pronounced sloshing motion. These conclusions hold for fluids or gases (where they are obvious) as well as for solids. As a practical conclusion it should be noted that the measurement of the imaginary parts of admittances poses some difficulties which should be taken care of when results are compared.

Fig.6 shows the admittances of solid layers of different thickness and frequency. In the one dimensional case; i.e. when a strip of width D is excited, the ordinate is

$$\tilde{A}_S = \frac{\hat{v}_3}{\hat{F}} \frac{G}{\omega} \tag{73}$$

in the two dimensional case; i.e. when a circle of diameter D is excited the ordinate is

$$\tilde{A}_P = \frac{\hat{v}_3}{\hat{F}} \frac{G}{\omega k_T} . \tag{73a}$$

The abzisse is $H = 2h\pi/\lambda_T = k_T h$ where h is <u>half</u> the thickness.

The curves are surprisingly smooth in spite of the fact that there are many cut-off frequencies. As has been expected the real parts are independent on source size whereas the imaginary parts change considerably; for the one dimensional case they have a logarithmic singularity as D goes to zero, for the two dimensional case there is a 1/D singularity. The limiting cases which follow either from the simple bending wave theory or from the calculations on the semi-infinite model are

$$\left.\begin{aligned} \mathrm{Re}\{\tilde{A}_S\} &= 0{,}14\sqrt[4]{(1-\nu)}\, H^{-1{,}5} \\ \mathrm{Im}\{\tilde{A}_S\} &= -0{,}14\sqrt[4]{1-\nu}\, H^{-1{,}5} \\ \mathrm{Re}\{\tilde{A}_P\} &= 0{,}077\sqrt{(1-\nu)}\, H^{-2} \end{aligned}\right\} H<1 \tag{74}$$

$$\left.\begin{aligned} \mathrm{Re}\{\tilde{A}_S\} &\approx 0{,}463(1-\nu) \\ \mathrm{Im}\{\tilde{A}_S\} &\approx 0{,}243(1-\nu)\ln\left[\{1.9-15(\nu-0{,}25)^2\}\lambda_T/D\right] \\ \mathrm{Re}\{\tilde{A}_P\} &\approx 0{,}19(1-\nu) \\ \mathrm{Im}\{\tilde{A}_P\} &\approx 0{,}09(1-\nu)\lambda_T/D . \end{aligned}\right\} H>1 \tag{75}$$

If one is satisfied with a 25% accuracy the real parts of the admittances can be estimated the following way:

a. determine the relevant wave length λ_{rel}; i.e. bending wave length for plates, Rayleigh wavelength for a very thick structure,

b. make a sphere of radius $\lambda_{rel}/4$ around the excitation (in the case of strip excitation one obviously has to make a circle of radius $\lambda_{rel}/4$,

c. determine the mass inside this sphere (or circle) and call it m_{rel},

d. calculate

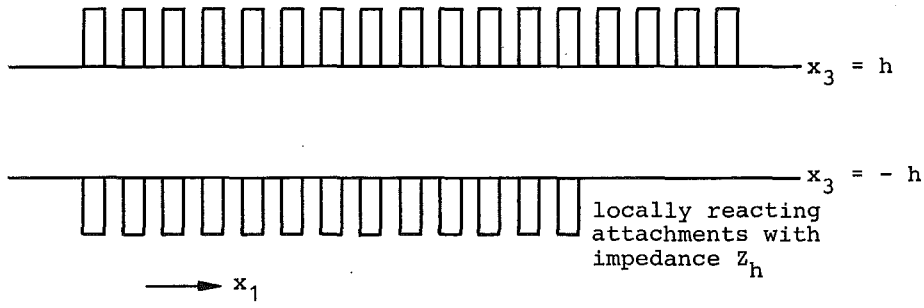
$$\mathrm{Re}\{A\} \approx \frac{1}{\omega m_{rel}} . \tag{76}$$

It easily can be checked that this formula holds not too badly when it is applied to a force acting on a fluid medium (dipole-radiator). Thus we have as a rather general rule: The real part of the admittance is roughly determined by the mass within the range $\lambda_{rel}/4$.

7. Wave guides with locally reacting absorbers

Lined ducts are very successfully used to reduce the sound transmission in air conditioning systems and other gas or fluid filled pipe configurations. The theoretical model which is used to describe the behaviour of such devices is rather simple; it consists of a wave guide with boundaries that have a given locally reacting impedance

$$Z_h = p(x_3 = \pm h) / v_3(x_3 = \pm h). \tag{77}$$



Using the same notation as in the previous sections but letting $k_2 = 0$ the boundary conditions at the planes $x_3 = \pm h$ are /5/,/6/

$$p_A e^{-jq_s h} + p_R e^{jq_s h} = Z_h \left(\frac{q_s}{\omega \bar{\rho}} p_A e^{-jq_s h} - \frac{q_s}{\omega \bar{\rho}} p_R e^{jq_s h} \right)$$

$$p_A e^{jq_s h} + p_R e^{-jq_s h} = Z_h \left(\frac{-q_s}{\omega \bar{\rho}} e^{jq_s h} + \frac{q_s}{\omega \bar{\rho}} \frac{p_R}{p_A} e^{-jq_s h} \right) p_A . \tag{78}$$

This system of linear equations with the unknowns p_A and p_R has a solution only if the determinant vanishes. This way the real and imaginary parts of q_s are given (as a transcendental equation); they in turn determine the wave number in the x_1-direction whose imaginary part is the decay constant for sound waves in the direction of the duct axis.

Thus it is possible to calculate the decay of sound in ducts if the impedances are given /6/,/7/. This is done quite often in practice and the fact that it is possible nowadays to build silencers that show a decay of more than 3 dB per duct height is to a large degree due to the fact that the sound propagation in such ducts is so well understood.

The situation is much less favorable for solid wave guides. In this case the theory is just at the beginning /8/ and the decay rates that are achieved in practice are much lower than those that are quite common in air borne sound problems in spite of the fact that the wave lengths in both cases are of the same order of magnitude.

This is somewhat surprising because the theoretical model for a solid wave guide with attachments is practically the same as the one used in the first part of this section to describe sound in gas or fluid filled wave guides; the only difference is that the equations are approximately twice as long.

If the figure between eq.(77) and (78) is understood as a solid wave guide and if it is assumed that the attachments do not generate shear forces the equations corresponding to eq.(78) are

$$\sigma_{31}(x_3 = h) = 0$$

$$\sigma_{31}(x_3 = -h) = 0$$

$$\sigma_{33Ac}\,e^{-jq_c h} + \sigma_{33Rc}\,e^{jq_c h} + \sigma_{33AT}\,e^{-jq_T h} + \sigma_{33RT}\,e^{jq_T h} =$$

$$= j\omega Z_h \left[\xi_{3Ac}\,e^{-jq_c h} + \xi_{3Rc}\,e^{jq_c h} + \xi_{3AT}\,e^{-jq_T h} + \xi_{3RT}\,e^{jq_T h} \right],$$

$$\sigma_{33Ac}\,e^{jq_c h} + \sigma_{33Rc}\,e^{-jq_c h} + \sigma_{33AT}\,e^{jq_T h} + \sigma_{33RT}\,e^{-jq_T h} =$$

$$= j\omega(-Z_h) \left[\xi_{3Ac}\,e^{jq_c h} + \xi_{3Rc}\,e^{-jq_c h} + \xi_{3AT}\,e^{jq_T h} + \xi_{3RT}\,e^{-jq_T h} \right], \qquad (79)$$

If σ_{33} and σ_{31} are expressed in terms of ξ_3 by using eq.(53)-(55) a set of four linear equations is obtained. Here again the vanishing of the determinant yields the free wave number and thus also the decay rate in x_1-direction.

If realistic impedances are introduced into equation (79) it turns out that at least in narrow frequency bands the decay rates can be extremely high, making it rather likely that effective "mufflers" for structural vibrations on beams and other solid wave guides can be built.

References

/1/ Morse, P.M.; Feshbach, H.: Methods of Theoretical Physics. Part I, Chapt. 1.6 McGraw - Hill Book Company, New York, (1953)

/2/ Lighthill, M.J.: On sound generated aerodynamically. I General Theory. Proc. Roy. Soc. A 211, 564-587 (1952)

/3/ Achenbach, J.D.: Wave propagation in elastic solids. Chapt. 2.4 North-Holland Publishing Company, Amsterdam (1973)

/4/ Cremer, L.; Heckl, M.; Ungar, E.E.: Structure-borne Sound. Chapt. II, 6, a Springer, Berlin (1975)

/5/ Morse, P.M.: The transmission of sound inside pipes. J. Acoust. Soc. Amer. 11, 205 (1939)

/6/ Cremer, L.: Theorie der Luftschalldämpfung im Rechteckkanal mit schluckender Wand und das sich dabei ergebende höchste Dämpfungsmaß. Acustica 3, 249 (1953)

/7/ Mechel, F.: Schalldämpfer. Chapt. 19 in Taschenbuch der Technischen Akustik Springer, Berlin (1973)

/8/ Heckl, M.: Körperschalldämpfung bei Balken durch seitlich angebrachte Widerstände. Acustica 45, 201 (1980)

/9/ Heckl, M.: Körperschallübertragung bei homogenen Platten beliebiger Dicke. To be published in Acustica 1981

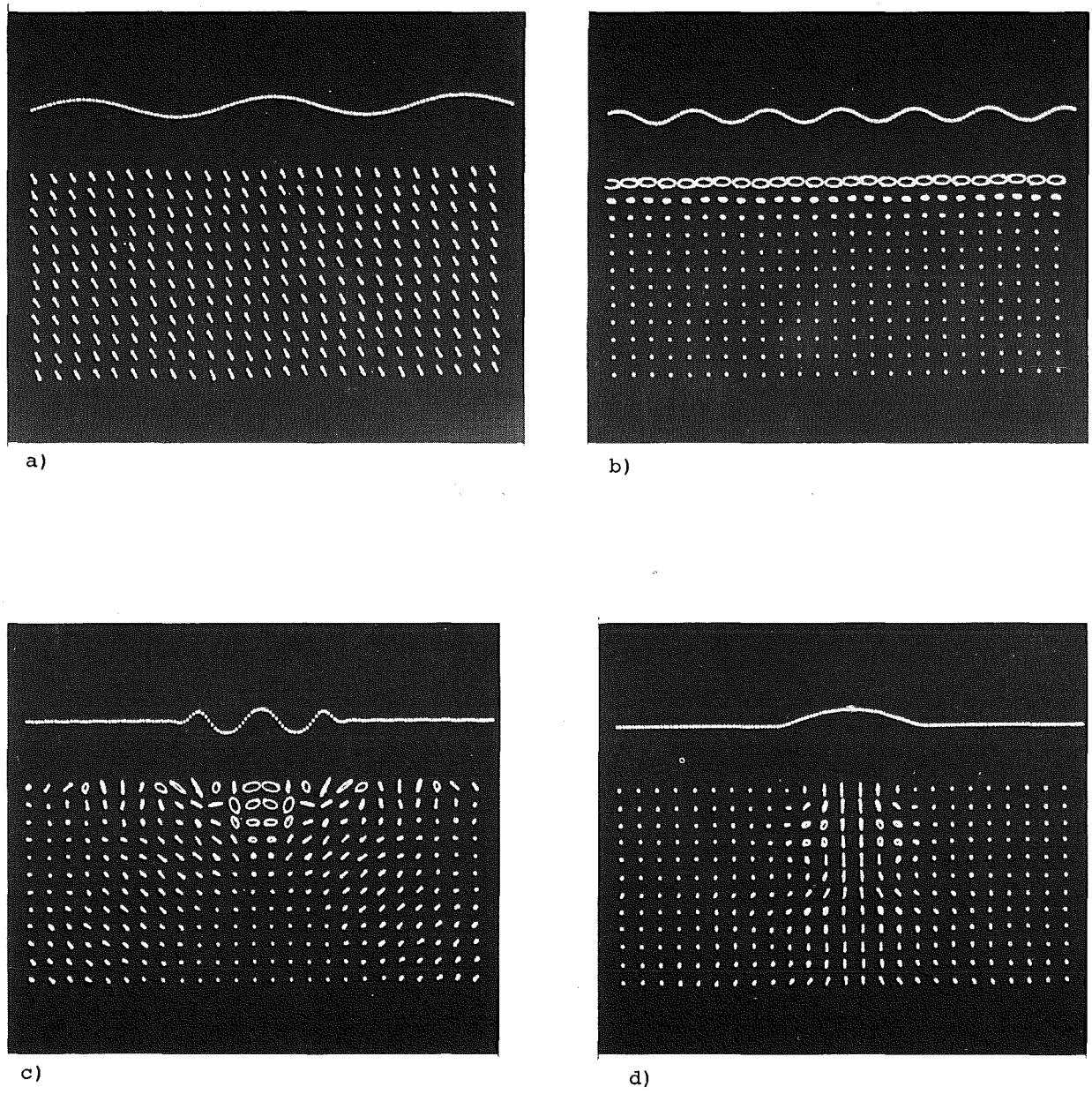


Fig. 1
Examples of calculated particle paths in fluids which are excited at a free surface. The upper line gives the spatial distribution of the excitation. (λ_e= wave length of the excitation).

a) $\lambda_e > \lambda_S$ infinite excitation area, above coincidence

b) $\lambda_e < \lambda_S$ infinite excitation area, below coincidince

c) $\lambda_e > \lambda_S$ finite excitation area.

d) plane radiator with constant phase.

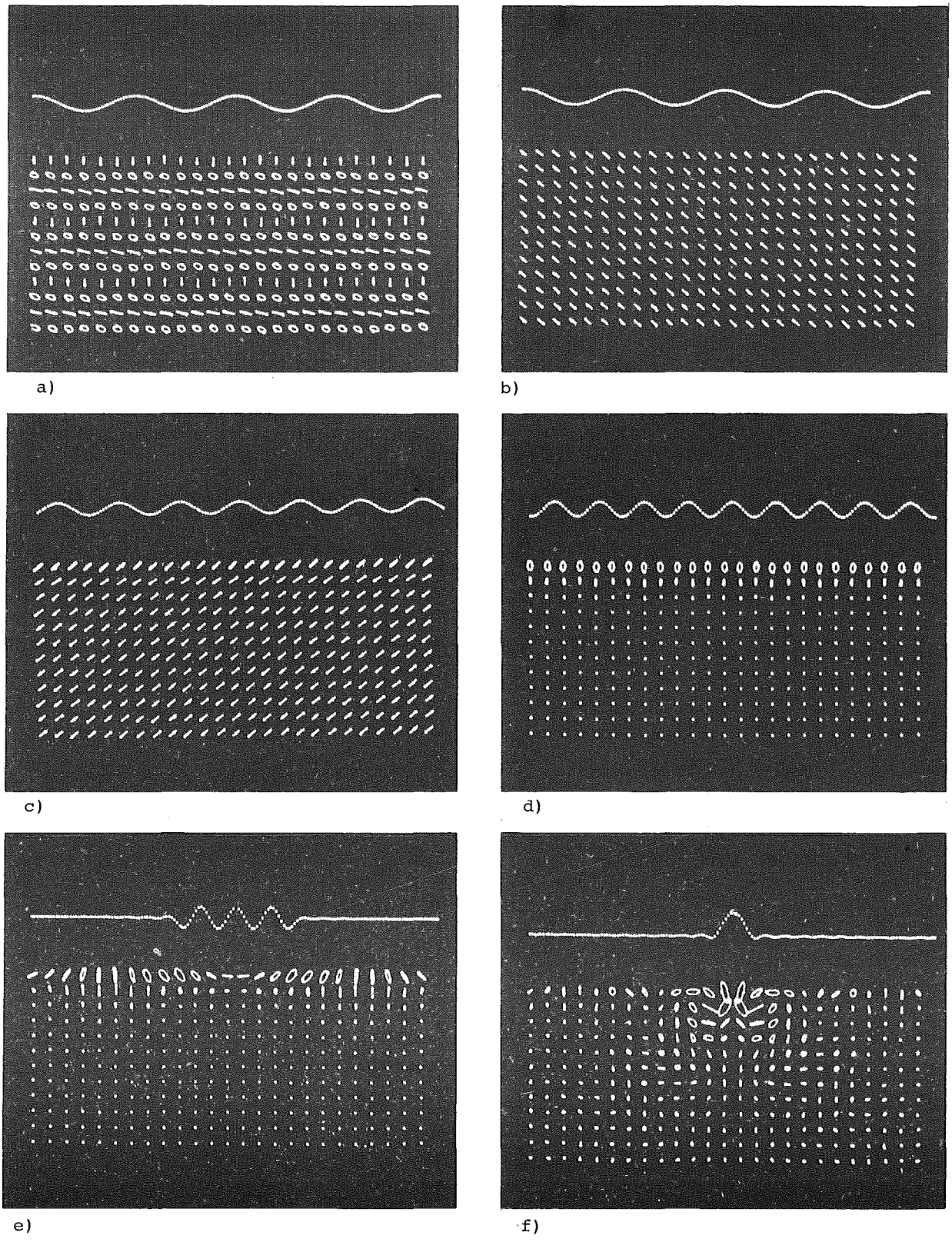


Fig. 2
Examples of calculated particle paths in solids which are excited at a free surface. The upper line gives the spatial distribution of the excitation.
a) $\lambda_e > \lambda_C > \lambda_T$ infinite excitation area. Compressional waves and shear waves are generated.
b) Some as a) but without shear waves.
c) $\lambda_C > \lambda_e > \lambda_T$ infinite excitation area. Mainly shear waves are generated.
d) $\lambda_C > \lambda_T > \lambda_e$ infinite excitation area. Only near field excitation.
e) $\lambda_C > \lambda_T > \lambda_e$ finite excitation area. Mainly Rayleigh waves.
f) Localised excitation.

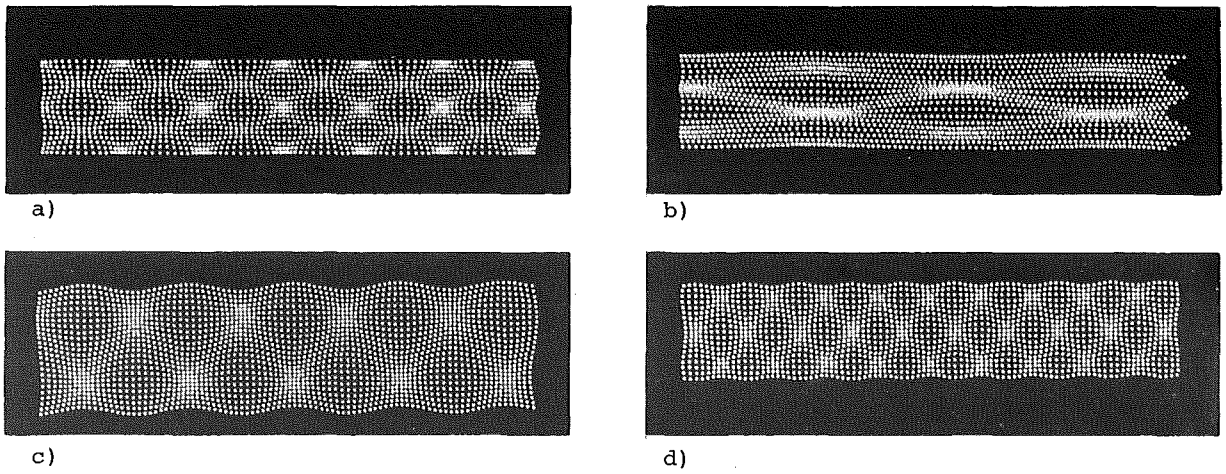


Fig. 3
Wave fields in fluid or gas filled ducts
a) and b) hard walls
c) and d) soft walls

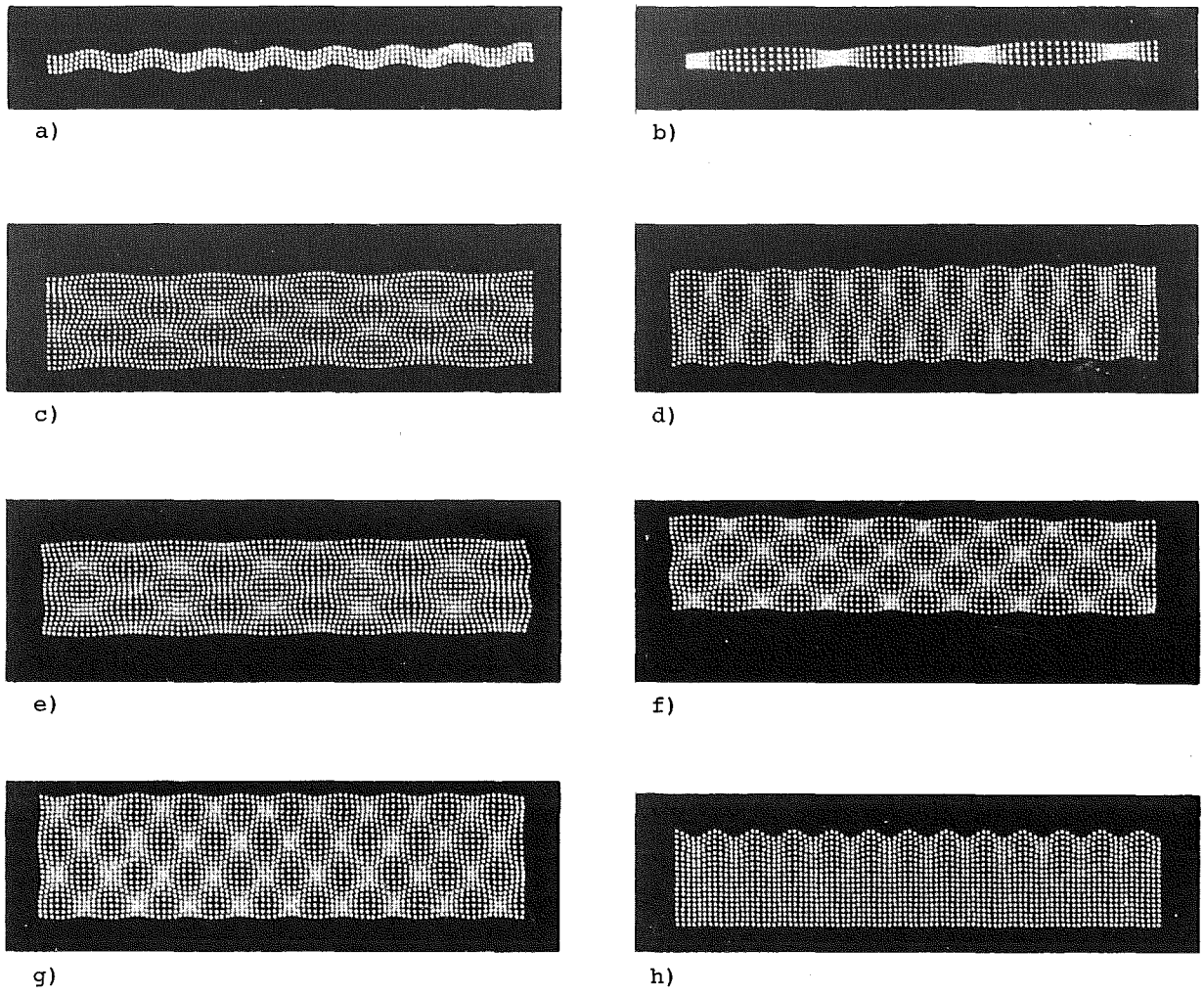


Fig. 4
Wave fields in solid ducts with free surface
a) bending wave
b) (quasi-) longitudinal wave
c)-h) higher order modes

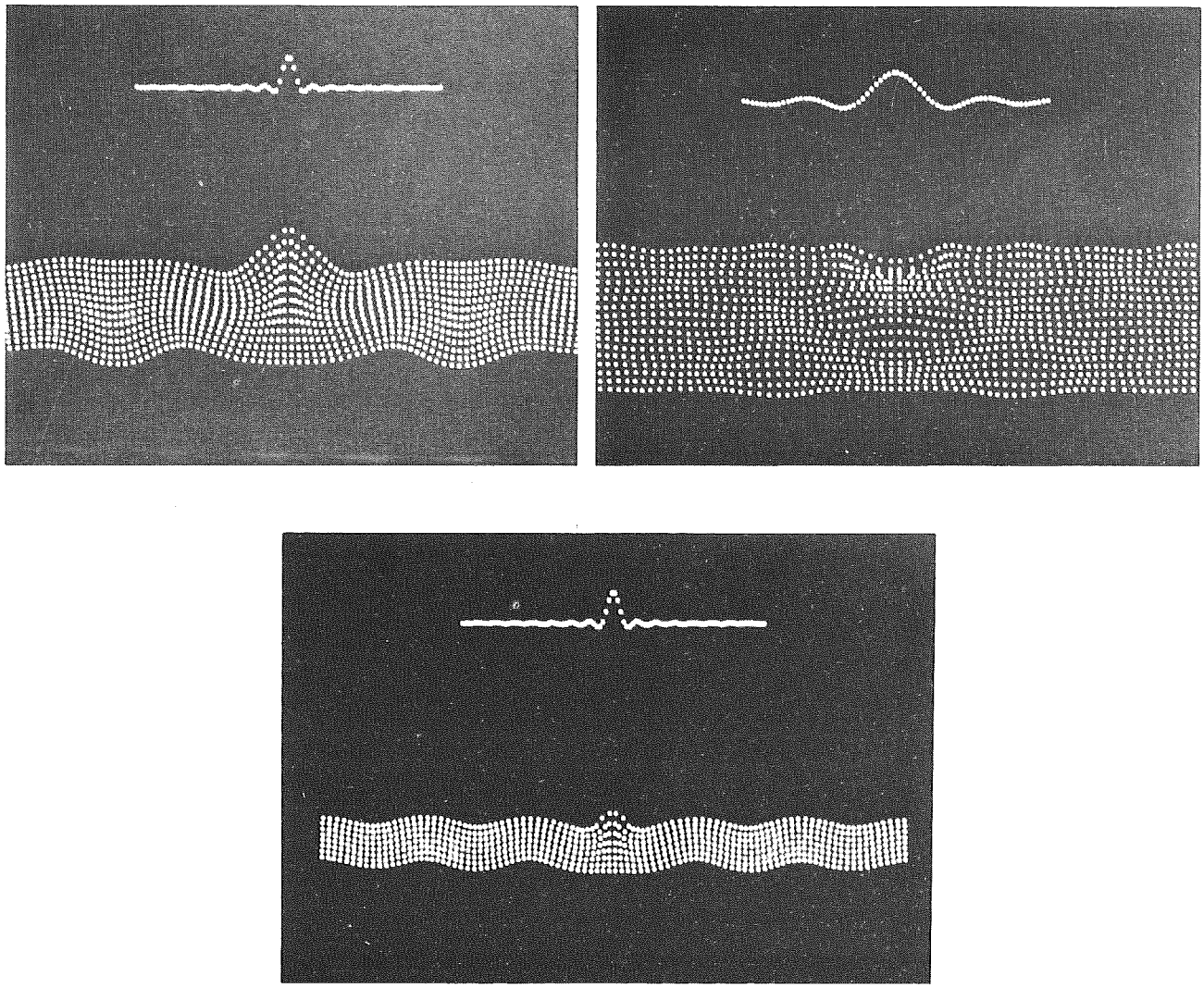


Fig. 5

Wave fields in thick plates that are excited by a localised pressure distribution (see upper line).

Continued from page 18, Fig. 6

B: Real part for excitation in a circle

curve a: Poisson's ratio = 0; $D/\lambda_T = 10^{-3}$

curve b: Poisson's ratio = 0,33; $D/\lambda_T = 10^{-3}$

curve c: Poisson's ratio = 0,5; $D/\lambda_T = 10^{-3}$

curve d: Theoretical value for a thin plate

C: Imaginary part for strip excitation

curve a: $D/\lambda_T = 0{,}316$

curve b: $D/\lambda_T = 0{,}1$

curve c: $D/\lambda_T = 0{,}0316$

curve d: $D/\lambda_T = 0{,}01$

curve e: $D/\lambda_T = 0{,}00316$

curve f: $D/\lambda_T = 0{,}001$

curve g: Theoretical value for thin beam.

D: Imaginary part for excitation in a circle

curve a: $D/\lambda_T = 0{,}316$

curve b: $D/\lambda_T = 0{,}1$

curve c: $D/\lambda_T = 0{,}0316$

curve d: $D/\lambda_T = 0{,}01$

curve e: $D/\lambda_T = 0{,}00316$

curve f: $D/\lambda_T = 0{,}001$

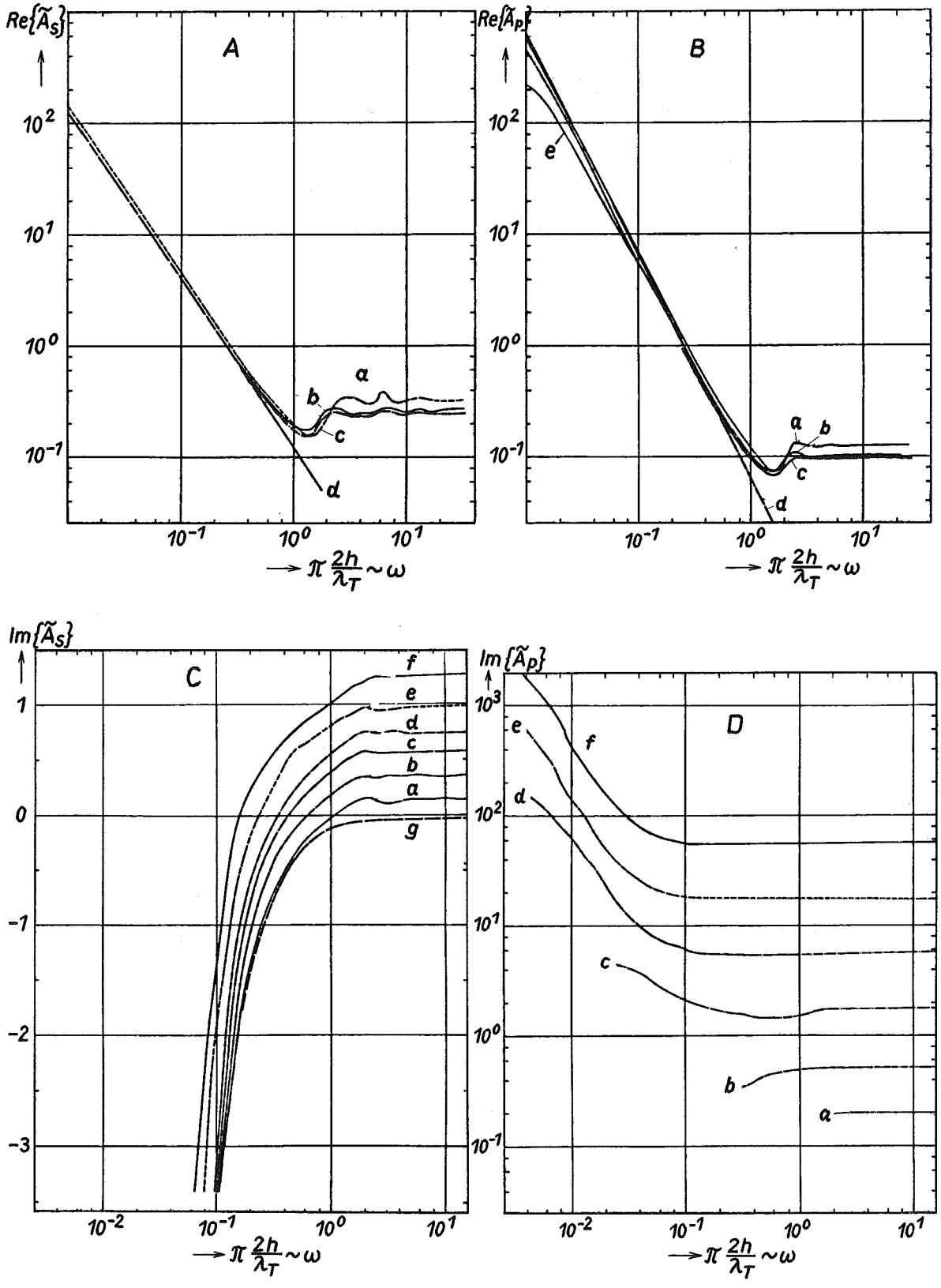


Fig. 6

Real and imaginary parts of the normalised admittance for excitation within a circle of diameter D (A_p) or a strip of width D (A_s).

A: Real part for strip excitation

curve a: Poisson's ratio = 0

curve b: Poisson's ratio = 0,33

curve c: Poisson's ratio = 0,5

curve d: Theoretical value for a thin beam

Continuation at page 17

PROCEDURES RELATING THE NEAR- TO THE FAR-FIELD: IMAGING TECHNIQUES

M. C. Junger
Cambridge Acoustical Associates, Inc.
54 Rindge Avenue Extension
Cambridge, Massachusetts 02140, USA

The formulation of the pressure field of an arbitrary sound source in terms of the Helmholtz integral is discussed. Analytical and experimental approaches to the evaluation of this integral are described. The latter emphasize the use of measurements taken in the near-field. The plane-wave asymptotic technique is described. The concept of intensity is introduced and its measurement for harmonic and random signals described. Finally, two types of computer holography are described: One of these provides a wavenumber filtering procedure which reconstructs only the supersonic components of the dynamic configuration of the source; the other reconstructs the actual configuration of the source; as well as the three-dimensional pressure and vector intensity fields.

The infinite homogeneously vibrating surface, Lecture 7, Section B.3, generates a pressure field uniquely related to its monochromatic line wavenumber spectrum. Specifically, if $k_s < k$, the train of waves radiates plane sound waves; if $k_s > k$, it generates incompressible near-field pressures. In other more realistic situations, the dynamic configuration of the radiating surface embodies a wavenumber spectrum resulting in a transform extending over all wavenumbers. This situation was encountered in the analysis of various sound sources: the piston, the finite cylinder, and the point-excited plate. In these situations, there does not exist a one-to-one relation between the dynamic configuration of the vibrating surface and the far-field pressure. Specifically, the transition from the vibrating surface or the near-field to the far-field is effectively a low-pass wavenumber filter which eliminates the subsonic portion $k < k_s$ of the wavenumber spectrum. In fact, vibrational patterns differing as to their subsonic spectrum generate similar far-fields if their supersonic wavenumber spectra coincide. There is therefore an incentive to determine which features of the structural vibrations are responsible for sound radiation. Several of these techniques will be described.

A. THE HELMHOLTZ INTEGRAL

Rayleigh's formula (Lecture 7, Section B.1) or the Hankel transform formulation (Lecture 7, Section B.2 to B.5) are confined to planar radiators located in a rigid baffle or endowed with a plane of symmetry. Other source configurations which allow the analytic construction of the far-field from information on their surface configuration are equally limited to certain geometries, e.g., the infinite cylinder. For arbitrary surfaces, the far-field is given in terms of the surface pressure as well as the normal accelerations of the radiating surface. An expression for the pressure, applicable both in the near- and far-field, is the Helmholtz Integral:[1]

$$p(\bar{R}) = \int_{S(\bar{R}_o)} \left(p \frac{\partial g}{\partial n} - g \frac{\partial p}{\partial n} \right) dS$$

where g is the free-space Green's function which satisfies the inhomogeneous Helmholtz equation

$$(\nabla^2+k^2) g(|\bar{R}-\bar{R}_o|) = -\delta(|\bar{R}-\bar{R}_o|)$$

The free-space Green's function is effectively the point source, Eq. 5.26, specialized to unit strength $\ddot{Q} \equiv 1$ and generalized to arbitrary source locations by replacing R by $|\bar{R}-\bar{R}_o|$. This function and its normal derivative are

$$g(|\bar{R}-\bar{R}_o|) = \frac{\exp(ik|\bar{R}-\bar{R}_o|)}{4\pi|\bar{R}-\bar{R}_o|} \tag{1}$$

$$\frac{\partial g}{\partial n} = \frac{g}{|\bar{R}-\bar{R}_o|} (1-ik|\bar{R}-\bar{R}_o|)\cos\theta$$

$$\simeq -ikg\cos\theta \quad , \quad k^2|\bar{R}-\bar{R}_o|^2 >> 1$$

where θ is the angle between the outward normal and the vector $\bar{R}-\bar{R}_o$.

The Helmholtz integral is a powerful tool which constructs a three-dimensional pressure field from the surface pressure and its normal derivative. This integral formulation is based on Gauss' Integral Theorem which relates volume and surface integrals. Introducing the boundary condition, Eq. 5.16, the Helmholtz integral can be written in terms of pressures and accelerations:

$$p(\bar{R}) = \int_S \left(p \frac{\partial g}{\partial n} + \rho \ddot{w} g \right) dS \tag{2}$$

On a planar surface, the pressure and acceleration integrals contribute equally. One can therefore drop the pressure component of the integral and multiply the acceleration integrand by 2. This yields Rayleigh's formula, Eq. 7.4. More generally, the surface pressure can be eliminated whenever there exists a Green's function

$$G = g+\Gamma \quad , \quad \frac{\partial G}{\partial n} = 0 \text{ on } S(\bar{R}_o) \tag{3}$$

where Γ is a solution of the homogeneous Helmholtz equation. If a Green's function of the form of Eq.3 can be constructed, the pressure component drops out and

$$p(\bar{R}) = \int_S G\rho\ddot{w}dS \tag{4}$$

This is effectively a monopole distribution. Alternatively, for these same source configurations, the far-field can be constructed from the surface pressures as a dipole distribution:

$$p(R) = \int_S p \frac{\partial G}{\partial n} dS \tag{5}$$

The analytical construction of Green's functions suitable for implementing either Eq. 4 or 5 is possible only for the few source configurations described by coordinate systems in which the wave equation is separable, viz. the infinite plane, the infinite cylinder, the sphere, the spheroid, and a few other shapes of little practical interest. For arbitrary geometries, if information is available on the normal acceleration only, a Fredholm integral equation in the surface pressure must be solved before one can evaluate the far-field by means of Eq. 2. This is achieved by asymptotic or computer techniques. Alternatively, both accelerations and surface pressures may in some situations be determined experimentally and used to approximate the integral in Eq. 2.

B. EXPERIMENTAL APPROACHES TO THE HELMHOLTZ INTEGRAL

1. The Direct Implementation and its Shortcomings

Measurements of surface pressure and acceleration over a grid of points can clearly be used as input for a finite-difference approximation to the Helmholtz integral, Eq. 2. Severe drawbacks of this "brute force" approach are that

a. phase must be monitored between all grid points simultaneously;

b. the spacing of grid points is determined by the structural wavenumber spectrum and is therefore laboriously fine for a structure with stiffening frames and other complexities;

c. information on surface pressures and accelerations associated with the subsonic portion of the wavenumber spectrum is discarded by phase cancellation when the integral is evaluated, thus making the procedure inherently inefficient and unnecessarily cumbersome.

An elementary example of the nature of this inefficiency is provided by a small source combining pulsating and reciprocating motions of the same amplitude. The former component radiates with monopole efficiency, its sound field being given by Eq. 5.26. The dipole pressure component is smaller by a factor kL, where L is the characteristic dimension of the source. Consequently, the dipole component is of no practical significance with regard to sound radiation, even though accelerometer measurements suggest an importance comparable to the monopole pulsation.

2. Plane-wave Asymptotic Technique

The Helmholtz integral expression, Eq. 2 is equally valid on a control surface surrounding the source. Selecting a surface at sufficient range to endow it with a radius of curvature large compared to the acoustic wavelength, the impedance $z(\bar{R})$ approaches the ρc-impedance as illustrated for the sphere in Eq. 5.28. Setting $z \simeq \rho c$, the surface pressure in the integral in Eq. 2 can be approximated as $\rho c\dot{w} = i\rho\ddot{w}/k$, provided the control surface coincides with the wave front. Alternatively, one can replace $\rho\ddot{w}$ with $-ik\dot{p}$. Another large-argument approximation consists in assuming that $k^2|\bar{R}-\bar{R}_o|^2 >> 1$ in the latter of Eqs. 1, thus permitting the omission of the unit term. Under these circumstances, the Helmholtz integral can be evaluated in terms of either acceleration or, more practically, pressure measurements:

$$p(\bar{R}) \simeq \rho \int_S g\ddot{w}(1+\cos\theta)dS$$

$$\simeq -ik \int_S gp(1+\cos\theta)dS$$

A second advantage of the plane-wave approach is that the grid spacing can be coarse, viz. somewhat less than the acoustic wavelength. A drawback this technique shares with the direct implementation of the Helmholtz integral discussed in the preceding section is that phase must be monitored simultaneously between measurement points. Discussions of the application of this technqiue to transducer array calibration are found in the literature.[2,3]

We now turn to two near-field techniques which recently benefited from considerable development work.

C. ACOUSTIC INTENSITY MEASUREMENTS

1. Definition

Sound intensity is a vector representing acoustic power flow in the direction of the fluid particle velocity. This power is normalized to the unit area of wave front:

$$\bar{I}(t) = p\bar{\dot{d}} \tag{6}$$

Consider first a pressure varying harmonically with time as $\cos\omega t$. The corresponding fluid particle velocity vector can be formally written in terms of a vector impedance $\bar{z}$:

$$\bar{\dot{d}} = p/\bar{z}$$

Like the pressure the particle velocity is harmonic but, except when the plane-wave prerequisits are met, it is phase-shifted:

$$\bar{\dot{d}}(\bar{R},t) = \bar{\dot{d}}(\bar{R})\cos(\omega t+\phi) \tag{6a}$$

The phase angle is determined by the ratio of the real and imaginary components of the impedance

$$\phi \equiv -\tan^{-1} x/r$$

The time-averaged intensity for harmonic time dependence is

$$\begin{aligned} \langle\bar{I}\rangle &= \frac{\omega}{2\pi}\int_0^{2\pi/\omega} p\bar{\dot{d}}\,dt \\ &= (p\bar{\dot{d}})_{rms}\cos\phi \end{aligned} \tag{7}$$

In the far-field, this reduces to

$$\langle I\rangle = p_{rms}^2/\rho c$$

where the vector notation is not required since the wave propagates radially outward in this region. The incentive to utilize the intensity concept as compared to the plane-wave asymptotic technique is that the former does not impose a minimum range or frequency. The direct implementation of the Helmholtz integral (Section B.1) offers this same advantage, but, like the plane-wave approach, it requires simultaneous measurement of phase at all grid points. The great advantage of the intensity concept is that only relative phase between two neighboring points need be measured. Furthermore, it yields information on the location of source areas responsible for power radiation. In contrast to other approaches, including computer holography (see next section), near-field intensity measurements do not permit construction of the far-field distribution-in-angle, precisely because the directivity pattern depends on the phase relations between grid points on the control surface. The picture is further confused by circulating power cells which do not contribute to the far-field.

2. Signal Processing in the Time Domain

About fifty years ago, Olson proposed a technique aimed at discriminating against incompressible "pseudo-sound" pressures by finite-difference measurement of intensity. This he accomplished by means of two probes spaced a distance D, small in terms of wavelengths. If p_a and p_b are the pressures measured by these two probes, then

$$p \simeq \frac{1}{2}(p_a+p_b) = \frac{1}{2}|p_a+p_b|\cos\omega t$$

and from Eq. 5.1 and Eq. 6a above:

$$\bar{\dot{d}} \simeq \frac{\hat{n}}{\rho ckD}|p_b-p_a|\cos(\omega t+\phi)$$

where the unit vector $\hat{n}$ identifies the direction where (p_a-p_b) is a maximum. Substituting these two relations in Eq. 7, one obtains time-averaged intensity limited however to harmonic signals:

$$\langle\bar{I}\rangle \simeq \frac{\hat{n}\cos\phi}{2\rho ckD}(p_{brms}^2 - p_{arms}^2) \tag{8}$$

For random signals, the velocity requires time integration of acceleration, i.e., of p_a-p_b. A further integration of the instantaneous intensity yields the time-average of this vector:

$$\langle\bar{I}\rangle = \frac{\hat{n}}{2D\rho T}\int_0^T\left[(p_a+p_b)\int_0^t (p_a-p_b)\,dt'\right]dt$$

For stationary signals, the product of a function by its time-derivative tends to zero if averaged over a long enough time interval. Consequently, only the two cross products in the above expression contribute to the time-averaged intensity:

$$\langle \bar{I} \rangle = \frac{\hat{n}}{D\rho T} \int_0^T \left(p_a \int_0^t p_b dt' \right) dt \tag{9}$$

3. Signal Processing in the Frequency Domain

Fahy[4] noted that the integral in Eq. 9 can be formulated in the frequency domain,

$$\langle \bar{I} \rangle = -\frac{\hat{n}}{2\pi D\rho} \int_0^\infty \frac{\mathrm{Im}G(p_a,p_b,f)}{f} df \tag{10}$$

where $f = \omega/2\pi$ and G is the one-sided cross spectral density. The evaluation of the cross spectral density is performed after digitization of the signal. This technique has the advantage that a circuit switching procedure can be used to eliminate instrument phase mismatch.

The time-averaged product in Eq. 6 can of course also be evaluated from a pressure measurement made near the radiating surface and an acceleration or velocity measurement on this surface.[5] The intensity component normal to the surface is

$$\langle I_n \rangle = \langle p\dot{w} \rangle$$

$$= \frac{-1}{2\pi} \int_0^\infty \frac{\mathrm{Im}G(p,\ddot{w},f)}{f} df$$

A comprehensive up-to-date description of intensity measurements tecniques can be found in the proceedings of a recent international congress.[6]

D. COMPUTER HOLOGRAPHY

This last technique circumvents some of the drawbacks of the approaches described above. It is, however, still at an early stage of development. Work performed to-date in the U.S., at the Pennsylvania State University indicates that computer holography can provide a detailed picture of the near- and far-field both with regard to pressure and vector intensity. It provides, furthermore, a means of filtering the wave-number spectrum of the vibrating surface. Two basic forms of the technique have been developed, both limited to date to planar radiating surfaces, even though, in principle, there does not appear to be an obstacle to their adaptation to convex sources.

1. Wavenumber Filtering

In both implementations, pressures are measured in a hologram plane parallel to the source plane. The latter coincides with the (z=0) plane, while the hologram plane is defined by $z=z_o$. In the first application, z_o is large enough to eliminate evanescent waves. In the notation of Eq. 9.14, $\exp(-k_s z_o) \ll 1$. Pressures $p(x_h,y_h,z_o)$ are measured by an array which must be sparse to avoid the formation of standing waves. The pressures on this control surface can be used to construct the 3-dimensional pressure field by means of Eq. 5. For a plane, G = 2g. One therefore doubles the derivative claculated in Eq. 1. The pressure in Eq. 5 becomes

$$p(\bar{R}) = -2ik \int_{S_h} pg\cos\theta \, dS \tag{11}$$

A result less restricted as to hologram plane range can be constructed by using the first, non-asymptotic expression given in Eq. 1 for $\partial g/\partial r$. The Green's function in Eq. 1 describes waves propagating outward into the semi-infinite space $z > z_o$. The Green's function can however also be used to reconstruct a given wave front at an earlier time $t-(z_o/c)$, viz. the wave front on the source plane S_s by the simple device of replacing k in the exponential in Eq. 1 by -k, thus restoring the phase of the sound waves as it leaves the source plane:[7,8]

$$p(x_s,y_s) = -2ik \int_{S_h} pg^*\cos\theta \, dS \tag{12}$$

The asterisk indicates the complex conjugate of Eq. 1, and where g^* is evaluated for the range from a point on the hologram plane to a point on the source plane:

$$|\bar{R}-\bar{R}_o| = [z_o^2+(x_h-x_s)^2+(y_h-y_s)^2]^{1/2}$$

$$\cos\theta = -\frac{z_o}{|\bar{R}-\bar{R}_o|} \tag{13}$$

The cosine is negative since $z_o > 0$ and hence $\theta > \pi/2$. The pressure on the source plane can be written explicitly as

$$p(x_s,y_s,0) = \frac{ik}{2\pi}\int_{-\infty}^{\infty} p(x_h,y_h,z_o)\ \frac{z_o \exp(-ik|\bar{R}-\bar{R}_o|)}{|\bar{R}-\bar{R}_o|^2}\ dx_h dy_h \tag{14}$$

This formulation implies a time dependance of the form exp(-iωt). If exp(jωt) is assumed, -i must be changed to +j. In theory, S_h should be infinite. In practice the integral has converged for an array measuring three times the linear dimension of the source plane.

When it is located in the far-field, the array senses, by definition, exclusively true sound pressures, rather than exponentially decaying "pseudo-sound". The reconstructed dynamic configuration of the source therefore does not embody subsonic structural wavenumbers. The dynamic configuration

$$\ddot{w}(x_s,y_s) = -\rho^{-1}\partial p/\partial z \quad , \quad z = 0$$

reconstructed from Eq. 14 is the simplest vibrational pattern capable of generating the measured near-field. In fact, uniqueness of the acoustic solution is not applicable to the subsonic portion of the wavenumber spectrum of the radiating surface.[7]

2. Reconstruction of the 3-Dimensional Field

More recently,[9,10] a reconstruction of the near-field has been attempted and successfully achieved. This requires an array located close to the radiating surface to insure that the evanescent waves, Eq. 9.14, are of sufficient amplitude to be detectable in the presence of true sound waves. The pseudo-sound amplifi-amplitude on the radiating surface is reconstructed by changing the sign of the real exponents characterizing evanescent waves, exp(-k_sz) being replaced by exp(k_sz). Consequently, the sign of all exponents, whether real or imaginary, is changed to restore the original phase of sound waves and amplitude of evanescent waves. The procedure can be described more rigorously by stating that ∂G/∂n is replaced by its reciprocal. The image thus reconstructed yields the surface pressure spectrum. The computer reconstruction is also used to calculate the distribution of ∂p/∂z, and hence the surface acceleration and the normal component of the intensity vector. These same quantities can of course be computed on other planes.

This technique was applied to a shaker-excited aluminum plate measuring 93 x 61 x 1.3 cm. The plate is stiffened with one frame. An array of 256 microphones covering a surface of 325 x 324 cm was used. Even though the elements were spaced 20 cm, a grid point distance of 5 cm was achieved by translating the array.

Figures 1 and 2 show the reconstructed normal intensity component on the plate. Circulating power cells, whereby a portion of the plate acts as a sink, are apparent, especially at the lower frequency. A drastic change in the distribution and amplitude of radiated sound results from the minor increase in frequency from Figures 1 to 2. The reconstruction of the z-component of the intensity field 61 cm from the plate (Fig. 3) shows considerable smoothing and the absence of power cells, which are therefore seen to be a near-field phenomenon. Figures 4 and 5 display the intensity vector in planes normal to the plate, the orientation of the plane being respectively parallel and perpendicular to the frame. The inability of near-field intensity measurements to yield information on far-field distribution-in-angle is clearly apparent from the change in directivity with range. The power cell in Fig. 5 is caused by a failure of the epoxy bond between plate and frame.

REFERENCES

1 H. Lamb, Hydrodynamics, 6th ed. (New York, Dover Publishing Co., 1945), pp. 498-499.

2 C. W. Horton and G. S. Innis, Jr., J. Acoust. Soc. Am. 33, 877 (1961).

3 D. D. Baker, J. Acoust. Soc. Am. 34, 1737 (1962).

4 F. J. Fahy, J. Acoust. Soc. Am. 62, 1057 (1977).

5 C. H. Hansen and D. A. Bies, J. Sound Vib. 62, 93 (1979).

6 International Congress on Recent Developments in Acoustic Intensity Measurements, Senlis, France, 30 Sept. - 2 Oct. 1981.

7 M. Strasberg and G. Maidanik, The Quantitative Significance of Images Reconstructed by Acoustical Holography, unpublished ms of paper QQ1 presented at a meeting of the Acoustical Soc. of Am. (15 June 1979).

8 J. Powers, Acoustical Holography, edited by L. W. Kessler (New York, Plenum, 1976), Vol. 7, pp. 193-205.

9 E. G. Williams, J. D. Maynard, and E. Skudrzyk, J. Acoust. Soc. Am. 68, 340 (1980).

10 E. G. Williams and J. D. Maynard, Phys. Rev. Letters 45, 554 (1980).

11 Figure made available by Dr. E. G. Williams.

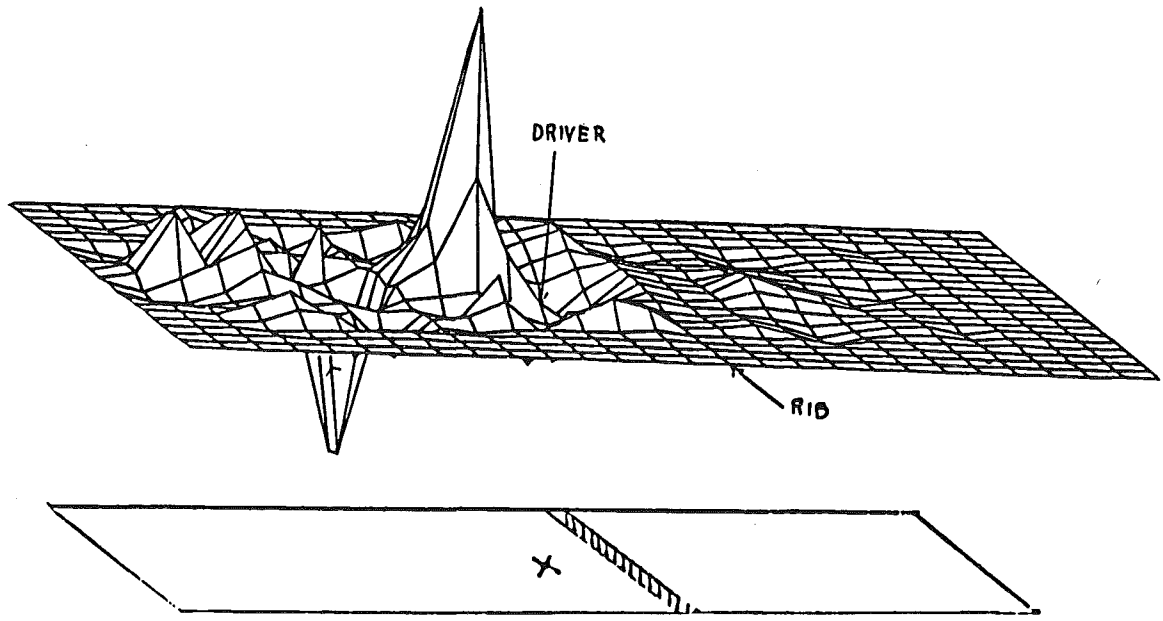


Fig. 1 - Normal component of the intensity vector on a frame-stiffened shaker-driven plate at 688 Hz.[11]

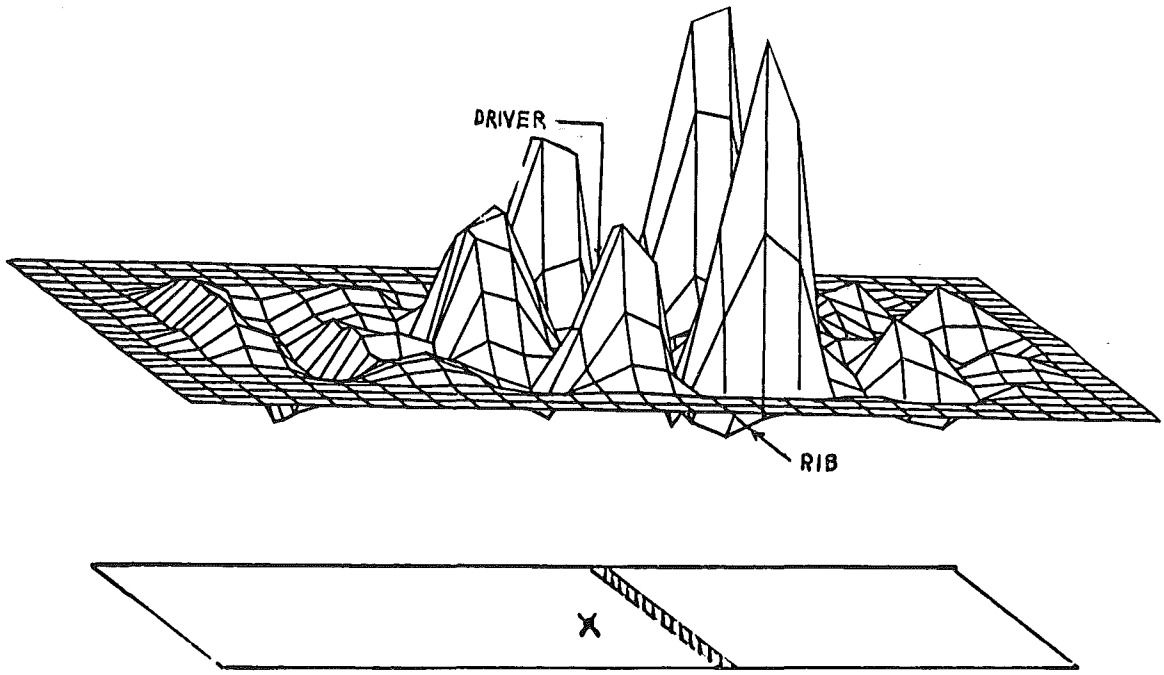


Fig. 2 - Normal component of the intensity vector on a frame-stiffened shaker-driven plate at 877 Hz.[11]

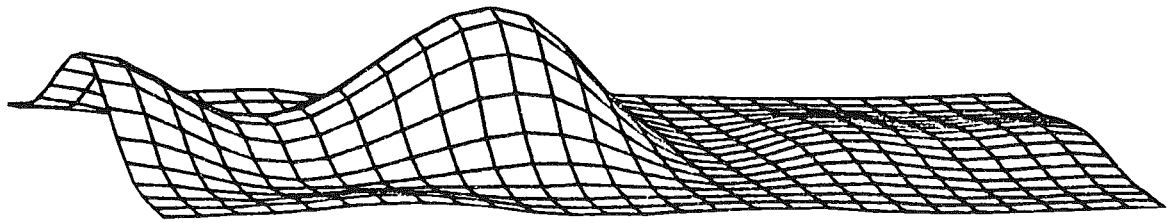


Fig. 3 - Normal component of the intensity vector on a frame-stiffened shaker-driven plate at 688 Hz on a plane located 61 cm from the plate.[11]

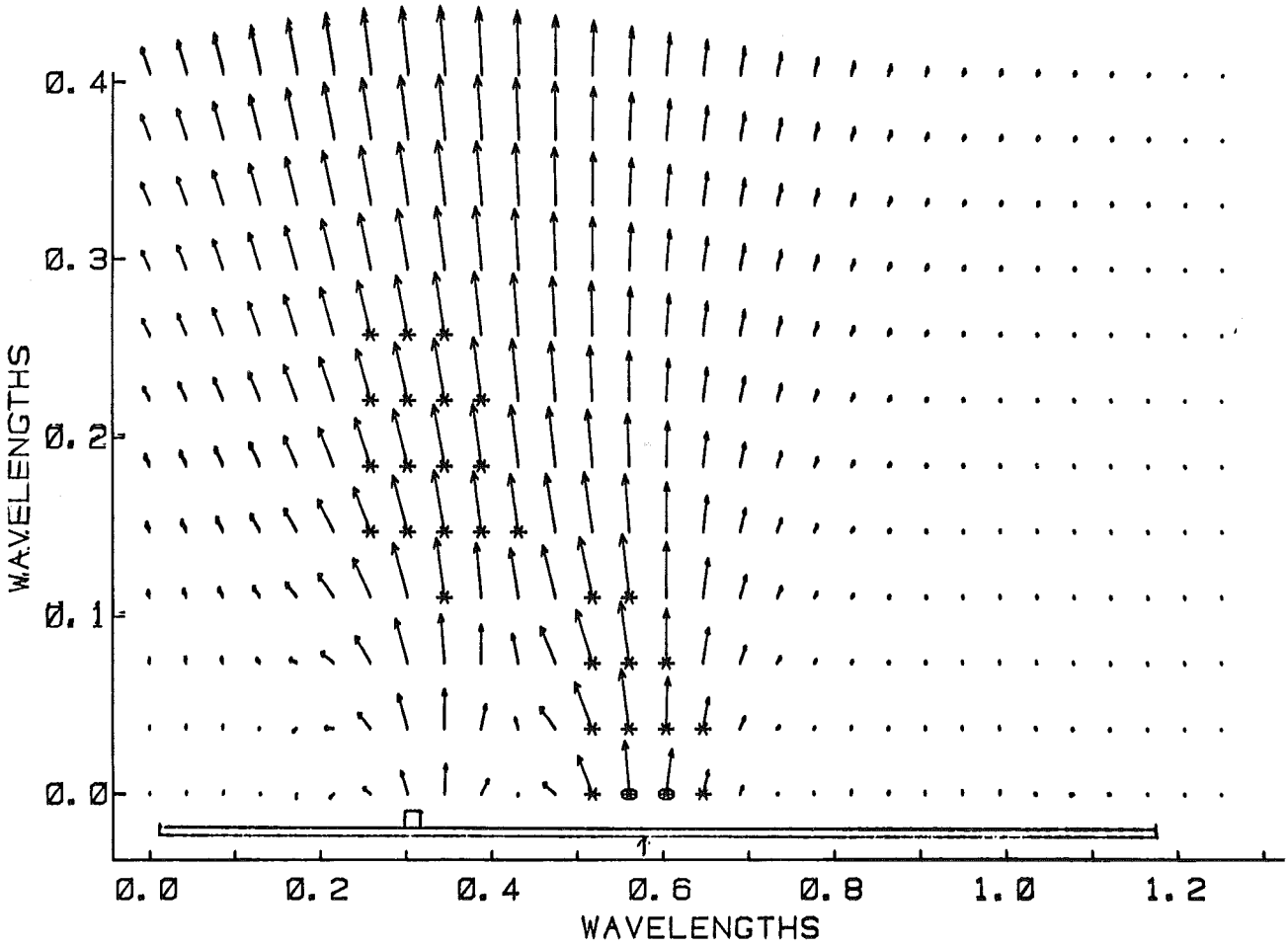


Fig. 4 - Intensity vector field in the plane of symmetry of the plate normal to the stiffening frame, at 330 Hz. The asterisk and the polygon identify vectors pointing respectively in and out of the paper. Vertical grid point spacing is 4.5 cm.[11]

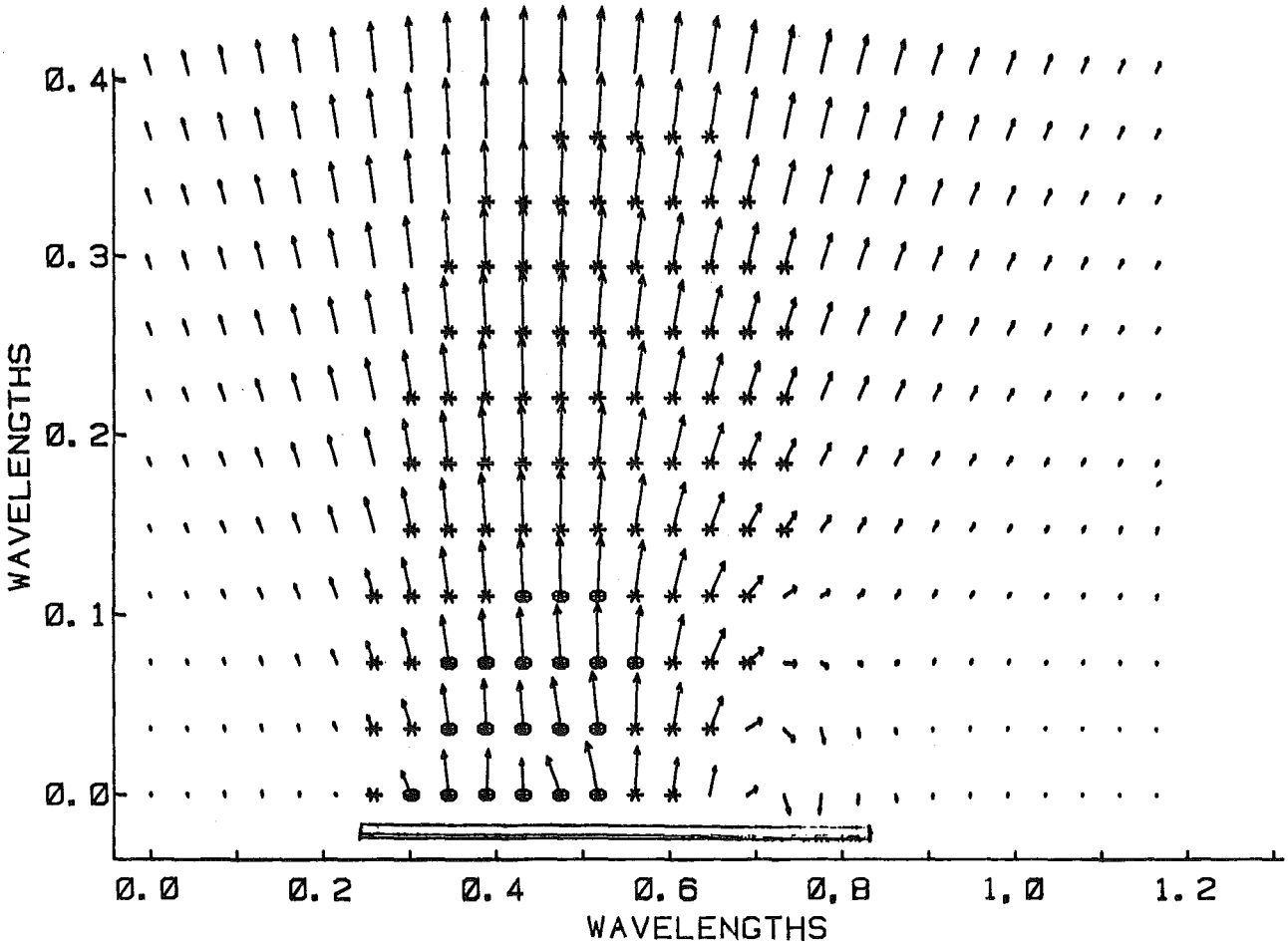


Fig. 5 - Intensity vector field in the plane containing the frame of the plate normal to the stiffening frame, at 330 Hz. The asterisk and the polygon identify vectors pointing respectively in and out of the paper. Vertical grid point spacing is 4.5 cm.[11]

PROGRES RECENTS DANS LA MESURE DE L'INTENSITE ACOUSTIQUE
RECENT DEVELOPMENTS IN ACOUSTIC INTENSITY MEASUREMENT

par M. PERULLI*
Chef de la Division Acoustique
Office National d'Etudes et de Recherches Aérospatiales, 92320 Châtillon - FRANCE

RESUME

Du 30 septembre au 02 octobre 1981, le Centre d'Etudes Techniques des Industries Mécaniques (CETIM) a organisé un Congrès International sur la mesure d'intensité acoustique.

Plus de 200 participants, représentant une vingtaine de nations des continents de l'Amérique du Nord, de l'Asie, de l'Australie et de l'Europe (de l'Est comme de l'Ouest), ont animé des débats de ce Congrès.

Compte tenu de la qualité des travaux exposés qui, véritablement, permettent de présenter l'état des connaissances dans ce domaine, en mettant en évidence les difficultés de mise en oeuvre, les réalisations actuelles et les développements futurs, une tentative de synthèse de ce Colloque sera exposée. Les légitimes contraintes d'édition ne permettent pas de fournir en temps utile le texte de cette synthèse qui sera remis aux participants au début de ce cours.

---

* et Professeur à l'Université de Technologie de Compiègne.

MODAL ANALYSIS USING DIGITAL TEST EQUIPMENT

by

Patrick Garcia
Hewlett-Packard France
Avenue Tropique
91 Les Ulis, France

SUMMARY

Modal Analysis methods based upon the measurement and post test processing of transfer functions in digital form are discussed. The analysis which shows how modal data can be identified from transfer function measurements is reviewed. This is followed by a discussion of alternative method for identifying modal data from transfer function measurements.

Finally, the effect of vibration in terms of acoustics is covered.

1. INTRODUCTION

At the present, the common use of the results of a modal analysis carried out on a mechanical structure consists of a decomposition of vibration into animated representation of the different mode shapes. In this way already a number of vibration as well as acoustic problems can be solved. Measures may be proposed based upon such an observation to get rid of some disturbing fatigue to stiffen the structure dynamically or to reduce the emitted noise.

There are a few techniques to analyse the dynamic behaviour of a structure due to vibration :

- equation of motion (ref.1) for a simple rigid body
- wave equation (ref.1) for an elastic, symetric body
- finite element method (ref.1) for complex structure
- and experimental methods as normal mode (ref.1) and transfer function approach.

In this paper, modal analysis based upon measurement and post test processing of the transfer function data in digital form are discussed. Since the FFT analyser provides a broad band frequency spectrum in real time, it can be used to get modal data of a structure from a serie of transfer functions.

Furthermore, if the input and response time signals are measured simultaneously, FFT transformed, and the transform of the response is divided by the transform of the input, a transfer function between input and response points on the structure is obtained. A FFT analyser is able to perform this operation quickly and accurately.

Modal analysis is the process of characterizing the dynamic behaviour properties of an elastic structure by identifying its modes of vibration. A mode of vibration is a global property of an elastic structure. It can be expressed in terms of resonance frequency, modal damping and mode shapes. Resonance frequency and modal damping can be identified from any point of the structure. The mode shape gives a spatial representation of the structure at the specified resonance frequency.

Modal testing is performed on mechanical structures in an effort to learn more about their elastic behaviour. Once these dynamic properties of the structures are known, its behaviour can be predicted, and therefore controlled and corrected on a vibration and acoustical point of view.

All these modal parameters can be used directly by a mechanical designer to modify the structure by adding mass, stiffness or damping in order to minimize the modes of vibration whose radiation efficiencies are important.

In this paper, the measurement of transfer functions in digital form and then the application of digital parameter identification techniques to identify modal parameters from the measured transfer function data are discussed. It is first shown that the transfer matrix which is a complete dynamic model of an elastic model of an elastic structure can be written in terms of the structure's modes of vibration. This special mathematical form allows one to identify the complete dynamics of the structure from a reduced set of test data, and is the essence of the modal approach to identifying the dynamics of a structure.

Next various considerations which apply to broad band testing using FFT analyser are discussed. Ways in which spectral resolution, measurement noise, can be handled with a digital machine are also discussed. Then the application of transfer function models and identification techniques for obtaining the modal parameters, single degree of freedom and multi-degree of freedom methods, are covered.

Finally one needs to know the acoustic radiation from the vibration structure.

2. THE STRUCTURE DYNAMIC MODEL

In the majority of present days modal analysis practice, the motion of the physical system is assumed to be adequately described by a set of simultaneous second order linear differential equation of the form

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = f(t) \quad (1)$$

where $f(t)$ = applied force vector
$x(t)$ = resulting displacement vector
$\dot{x}(t)$ = resulting velocity vector
$\ddot{x}(t)$ = resulting acceleration vector

and M, C, K are called the mass, damping and stiffness matrices. If the system has n dimension or n degrees of freedom, then the above vectors are n dimensional and the matrices are n x n. The mass, stiffness and damping matrices are usually assumed to be real valued and symmetric.
If the damping matrix is either proportional to the mass, and stiffness matrices (proportional damping), or can be replaced by a complex valued stiffness matrix (structural damping),or does not exist at all, the modal analysis is called normal mode practice.

In this approach, the modal parameters are identified by transfer function measurements. Therefore it is convenient to write equations (1) in their equivalent transfer function form.
Taking the Laplace transform of the system equation (1) gives :

$$B(s)\,X(s) = F(s) \qquad (2)$$

where F(s) : Laplace transform of applied force vector.
X(s) : Laplace transform of displacement vector.

$$B(s) = Ms^2 + Cs + K$$

B(s) is referred as the system matrix, and $B(s)^{-1}$ as the transfer matrix H(s) :

$$H(s) = B(s)^{-1} \qquad (3)$$

Hence H(s) satisfies the following equation :

$$X(s) = H(s) \cdot F(s) \qquad (4)$$

For a n.dimensional system, the matrix H(s) can be written :

$$H(s) = \begin{bmatrix} h_{11}(s) \ldots & h_{1n}(s) \\ \vdots & \\ & h_{ij}(s) \\ \vdots & \\ h_{n1}(s) & h_{nn}(s) \end{bmatrix} \qquad (5)$$

Where hij(s) is the transfer function in the i^{th} row and j^{th} column of the transfer matrix.

Typical transfer functions of one single degree of freedom is showed in figure 1.

It will be shown in the next section that it is possible to identify all modal parameters from measured transfer functions to specify the dynamic characteritics of the structure.

3. MODAL PARAMETERS : RESONANCE FREQUENCY, DAMPING AND MODE SHAPES.

The modes of vibration are defined in mathematical terms and it is shown that the transfer matrix can be written in terms of modal frequency, damping and modal vectors.

Since the transfer matrix H(s) is defined as the inverse of the system matrix B(s), it follows that the elements of H(s) are ratio of polynomials in s, with the determinant of B(s) in each denominator.

The transfer function hij can be written :

$$h_{ij}(s) = \frac{b_1 s^{2n.2} + b_2 s^{2n.1} \ldots\ldots + b_{2n-1} s + b_{2n-2}}{\det [B(s)]} \qquad (6)$$

where the b's are polynomial coefficients.
If the system is n.dimensional then det $[B(s)]$ is always a polinomial of order 2n roots, i.e values of s for which det $[B(s)] = 0$

If the roots of det $[B(s)]$ are assumed to be distinct, then H(s) can always be written in the partial fraction form

$$H(s) = \sum_{k=1}^{2n} \frac{A_k}{S - P_k} \qquad (7)$$

where P_k : k^{th} root of det $[B(s)] = 0$

Ak : residue matrix for k^{th} root (nxn matrix)

the roots P_k are referred to as poles of the transfer function.

From figure [2], the system is said subcritically damped, when the poles are complex numbers. Otherwise the structure is critically or supercritically damped when the poles are real values or lie along the real axis of the s.plane.
Each complex conjugate pair of poles corresponds to a mode of vibration in the structure. They are complex numbers, written as

$$P_k = -\sigma_k + i w_k \qquad P_k^* = -\sigma_k - i w_k \tag{8}$$

σ_k is the modal damping coefficient and w_k, natural frequency.

Then the resonant frequency is given by

$$\Omega_k = \sqrt{\sigma_k^2 + w_k^2} \tag{9}$$

and damping factor or percent of critical damping are

$$\varepsilon_k = \frac{\sigma_k}{\Omega_k} \tag{10}$$

Therefore when $\varepsilon_k < 1$, the mode (k) is subcritically damped.

Modal vector (U_k) are defined as solutions to the homogeneous equation

$$B[P_k]\ U_k = 0$$

According to ref [2], the transfer matrix can be written

$$H(s) = \sum_{k=1}^{2n} \frac{a_k U_k U_k^t}{s - P_k} \tag{11}$$

a_k being a scaling constant.

Then the modal residue matrix A for mode (k) can be written in terms of the modal vectors.

$$A_k = a_k U_k U_k^t \tag{12}$$

Shown in ref [2], the transfer matrix can be written as a summation of (n) conjugate pairs :

$$H(s) = \sum_{k=1}^{n} \left[\frac{U_k U_k^t}{s - P_k} + \frac{U_k^* U_k^{*t}}{s - P_k^*} \right] \tag{13}$$

where U_k^* is the conjugate of U_k and P_k^* the conjugate of P_k.

Only one row or column of the transfer matrix needs to be measured in order to identify all modal parameters of a structure provided the following assumptions are met.

1. The motion is linear, described by the linear second order equations.
2. The symmetry of motion or reciprocity [i.e. B and H matrices are symmetric]
3. No more than one mode exists at each pole location of the system transfer matrix.

4. MODAL TESTING CONSIDERATIONS

In the last section, it was shown that only one row or one column of the transfer matrix need to be measured in order to specify the modal parameters characterizing the mode of vibration of the structure.
In this section, various method of measuring transfer functions with a digital machine are discussed.
In addition, digital techniques for reducing measurement noise and increasing frequency resolution are discussed.

1. Two channel measurement with broadband excitation.

   Sine wave excitation can be used to measure transfer function, but presents some disavantages compared

to other types of signals : this technique is slow and tends to invalidate the modal parameter estimates since a dynamic behaviour of the structure presents some non linearity. A broadband random signal as transient signal, random testing is actually helpful in removing the effects of non linear behaviour from the measured transfer function data.

Transient testing

As fast method of performing transient testing is to use a hand held hammer with a load cell attached to impact the structure. The load cell measures the input force, and accelerometers mounted on the structure measure the response signals. The set up time required to use this method is negligible compared to mounting and alignment of shakers. A typical test is shown in fig.(3).
One way to control the frequency band of excitation on the structure is to use hard or soft head placed on the hammer.

Random testing

Two primary advantages of using a shaker driven by a broadband random signal are that it is more controllable than a transient signal and when used in conjunction with power spectrum ensemble averaging, noise and non linear components of the structural dynamics can be removed from the measured data.

Method ≠ 1

A straight forward approach is to use a random signal generator and pass the signal through the band pass filter of interest. The measured signals are not periodic in the measurement time window. Therefore the frequency spectrum will contain so called "leakage" causing inaccurate result. The Hanning window is normally used to reduce this "leakage" effect in the spectrum of a non periodic random signal.

Method ≠ 2

To avoid the problems of using a non periodic random signal, it is possible with a digital computer to generate a random periodic signal and to output it to the shaker through a digital to analog converter (D.A.C.). This signal synthesized by the computer is referred as pseudo random signal. However, because it is periodic, a pseudo random signal will not effectively remove non linear distortion components from the measurements.

Method ≠ 3

The best possible random signal is one that is both periodic in the window, i.e. satisfies the conditions for a periodic signal and yet changes with time so short it excites the structure in a truly random manner. This type of signal is referred to as a periodic signal. Reference (2) explains these methods in more detail.

2. Reducing measurement noise

One of the problems of any modal testing system is that extraneous noise from various sources in the measurement equipment or on the structure itself. Since we are interested in identifying modal parameters from measured transfer functions, the variance on the parameter estimates is reduced in proportion to the amount of noise reduction in the measurements.

Two methods of noise reduction which take advantage of the digital data form. One is an averaging technique and the other is a smoothing technique which is applied after the measurements have been made.

Averaging technique : The first technique tends to accumulate frequency measurements with a certain number of averaging. Before being FFT transformed, the signal has been weighted by a Hanning window. In order to improve the statistical properties of the Fourier transformed signal, a overlapping processing technique is used.

Exponential smoothing : If after a reasonable number of averages, the transfer function measurements are still relatively noisy, the noise can be further removed by means of exponential smoothing process. The exponential weighting process has the advantage to give an effect on the data which is known. The impulse response of a system, or the inverse Fourier transform of a transfer function can be expressed as a sum of the impulse response of the modes, i.e.

$$x(t) = \sum_{k=1}^{n} |R_k| \, e^{-\sigma_k t} \sin(\omega_k t + \alpha_k) \qquad (14)$$

where

$|R_k|$ = magnitude of the residue

α_k = phase angle of the residue

or

$$\alpha_k = \text{arctang} \frac{\text{imag. part } R_k}{\text{real part } R_k}$$

σ_k = damping coefficient

ω_k = natural frequency

Since the noise is distributed evenly throughout the interval of time, and the signal decreases exponentially.

the signal to noise ratio of the combined signal decays exponentially. Now if the noisy impulse response is weighted by an exponential window the signal information is preserved, whereas the noise is limited. Transformed back to the frequency domain, the corresponding transfer function is smoothed. Nevertheless weighting artificially an impulse response corresponds to an known increasing of the damping coefficient of the transfer function. Usually this known amount of damping is taken into account.

3. Gaining Frequency Resolution

Normally, the Fourier transform is computed in a frequency range from zero frequency to some maximum frequency. The digital Fourier transform is spread over a fixed number of frequency lines (typically 1024) which therefore limits the frequency resolution between lines. Band selectable Fourier analysis (B.S.F.A.) the so called ZOOM transform is a measurement technique in which Fourier transform based on digital spectrum analysis is performed over a frequency band whose upper and lower frequencies are independendly selected. B.S.F.A. can provide an improvement in frequency resolution of more than a factor of 100, as well as 10 dB increase in dynamic range, compared to baseband Fourier Analysis. It is also called the ZOOM transform since it zooms in on a portion of a baseband spectrum and "magnifies" it with mode lines of definition much as a camera zoom lens magnifies a picture.

4. MODAL PARAMETER IDENTIFICATION

When a structure is excited by a broadband input force, many of its modes of vibration are excited simultaneously. Since the structure behaves linearly, the transfer function can be expressed as the sum of different one single degree of freedom systems. At a given frequency the transfer function represents the sum of the modes of vibration which have been excited for this particular frequency. However, near the natural frequency of a particular mode its contribution to the overall motion is generally the greatest. The degree of mode overlap, i.e. the contribution of adjacent modes of vibration, is governed by the amount of damping of the modes and the frequency separation.
In case where modal overlap is light, the transfer function data can be considered in the vicinity of each modal resonance as if it was a single degree of freedom system. In other words it is assumed that the contribution of the tails of adjacent modes near each modal resonance is negligibly small (ref. 3).
On the other hand when modal overlap is heavy a single degree of freedom to modal parameter identification will not work ; the parameters of a certain number of modes have to be identified altogether (ref. 4).

5. RADIATION FROM VIBRATING STRUCTURES

From a good knowledge of the dynamic behaviour of a structure, it is now possible to obtain the vibration energy of this structure by knowing the excitation forces. Otherwise one gets experimentally the total vibration energy directly from the structure by using transducers. The most important vibration resonant frequency does not mean obligatory that they have the most effective acoutic radiation.

First it is necessary to review shortly some principles of radiation efficiency, then some simple examples in order to link vibration effects and emitted noise. The sound radiation efficiency describes the relation between structural vibrations an acoustical radiated power. This quantity is usually designated as

$$R = \frac{W}{\rho c\, S\, v^2} \qquad (15)$$

where W represents the sound power radiated by an structure with surface area S, and v the average R.M.S. velocity of the radiating surface.
Some derivations can be made for other source configurations, only the results for an infinite plate, submitted to bending waves will be discussed here, since they are of interest with respect to the experiments carried out. It is showed that the radiation efficiency for an infinite plate with bending waves is given by

$$R = \frac{k}{k - k_s} \quad \text{for } k_s \ll k \quad \text{or} \quad \lambda_s \ll \lambda \qquad (16)$$

where k_s = wave number of the plate vibrations
λ_s = wave l enght of the plate vibrations

As indicated this expression is only valid if the wavelength of the plate vibrations is greater than the wavelength in the ambient medium (air). For the situations where the wavelength of the plate vibrations is considerably smaller than the wavelenght in the air ($k_s \gg k$) sound pressure is 90° out of phase with plate velocity and the sound decreases. The frequency which separates regions is the frequency where bending waves of the plate travel as fast as the acoustic waves in the air ($k = k_s$), this frequency is called the critical frequency.

After some computation, the velocity of propagation of a bending wave on an infinite plate is

$$c_s = \sqrt{1.8\, h\, f\, c'_\ell} \quad , \quad c'_\ell \simeq \sqrt{\frac{E}{\rho_p}} \qquad (17)$$

where h is the thickness of plate, f the frequency of the wave, E Young's modulus, and ρ_p density of the plate.
The wavelength of the bending wave in the plate is given by

$$\lambda_s = \frac{c_s}{f} = \sqrt{\frac{1.8\,h\,c_L}{f}} \qquad (18)$$

In the case of a bending wave, the wavelength is proportionnal to the square root of the frequency. Because bending waves of different frequencies travel of different velocities, the waveform of a complex wave is not preserved, and the medium when excited into bending waves is said to be dispersive. Then the critical frequency f_c is defined as the frequency where $\varphi_s = \beta$ or $c_s = c$. From equation one gets

$$f_c = \frac{c^2}{1.8\,h} \sqrt{\frac{\rho_p}{E}} \qquad (19)$$

The equation (19) is strictly valid only for case where the wavelength of the flexural wave is greater than six times the thickness h of the panel ($\lambda_s > 6h$).

Thus for a specified structure the first modes whose resonance frequencies are above the critical frequency fc, radiate sound whereas the modes shapes whose resonance frequencies are smaller than the critical frequencies will have a bad radiation efficiency.

6. CONCLUSION

The experimental modal analysis is based on a black box approach ; for n points defining the structure n transfer functions are performed, therefore only one column or one row of the transfer matrix is obtained, assuming some properties of linearity, reciprocity.

The choice of the adequate algorithm of curve fitting determines the accuracy of the modal parameters. It must depend on the type of structure. In the case of coupled modes, a mono mode curve fitting would not be able to identify each mode accurately, therefore a multi mode approach has to be used.

Once all modal parameters identified from the transfer function measurements, FFT analysis gives an animation of the mode shape at the different calculated modes. This performance makes easier to understand how the structure vibrates ; mode of torsion, of flexion... (fig. 4)

From all these information the dynamic behaviour of the structure, it is interesting to know how to modify this structure in order i.e. to shift some disturbing resonance frequencies, or to decrease the displacement due to vibration at a point of the structure, without retesting the structure (ref. 5). Some analytical models are already existing relating the sensibility of the resonance frequency, damping coefficient and mode shapes to the modification of mass, stiffness and damping.

Nevertheless the aim tending to minimize the effect of vibration does not necessarily correspond to the acoustical annoyance. Therefore some other factors as radiation efficiency have to be taken into account when the goal is to handle with acoustics.

REFERENCES

1. Mark RICHARDSON — "Measurement and Characteritics of Dynamic Structures." HEWLETT-PACKARD - Santa Clara.

2. P.R. ROTH — "Effective Measurements using Digital Signal Analysis" IEEE Spectrum. Pp 62, 70, April 1971.

3. KENNEDY and PANCU — "Use of vectors in vibration Measurement and Analysis" J. Aerospace, Sci. vol 14, No 11, 1947.

4. SPITZNOGGLE and QUAZI — "Representation and Analysis of time limited signals using a complex exponential algorithm" J.A.S.A. vol 47, n° 5, 1970.

5. P. VANHONACKER — "An introduction to sensitivity analysis." Seminar of Modal Analysis. University of Leuven.1981.

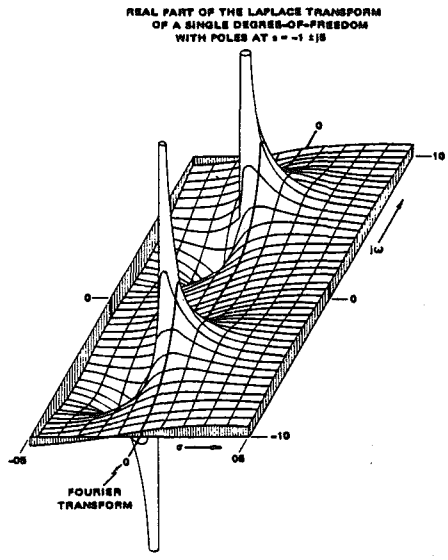


Figure 1. Real Part of a Transfer Function

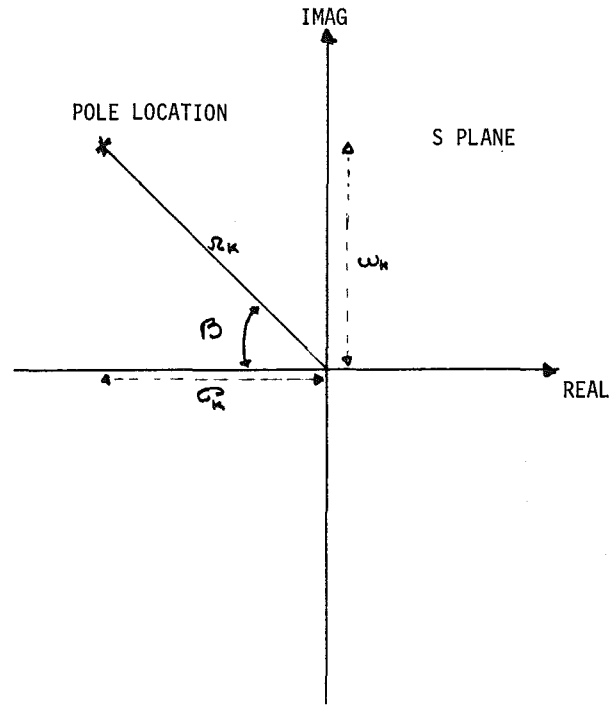


fig. 2 POLES OF A MODE (k)

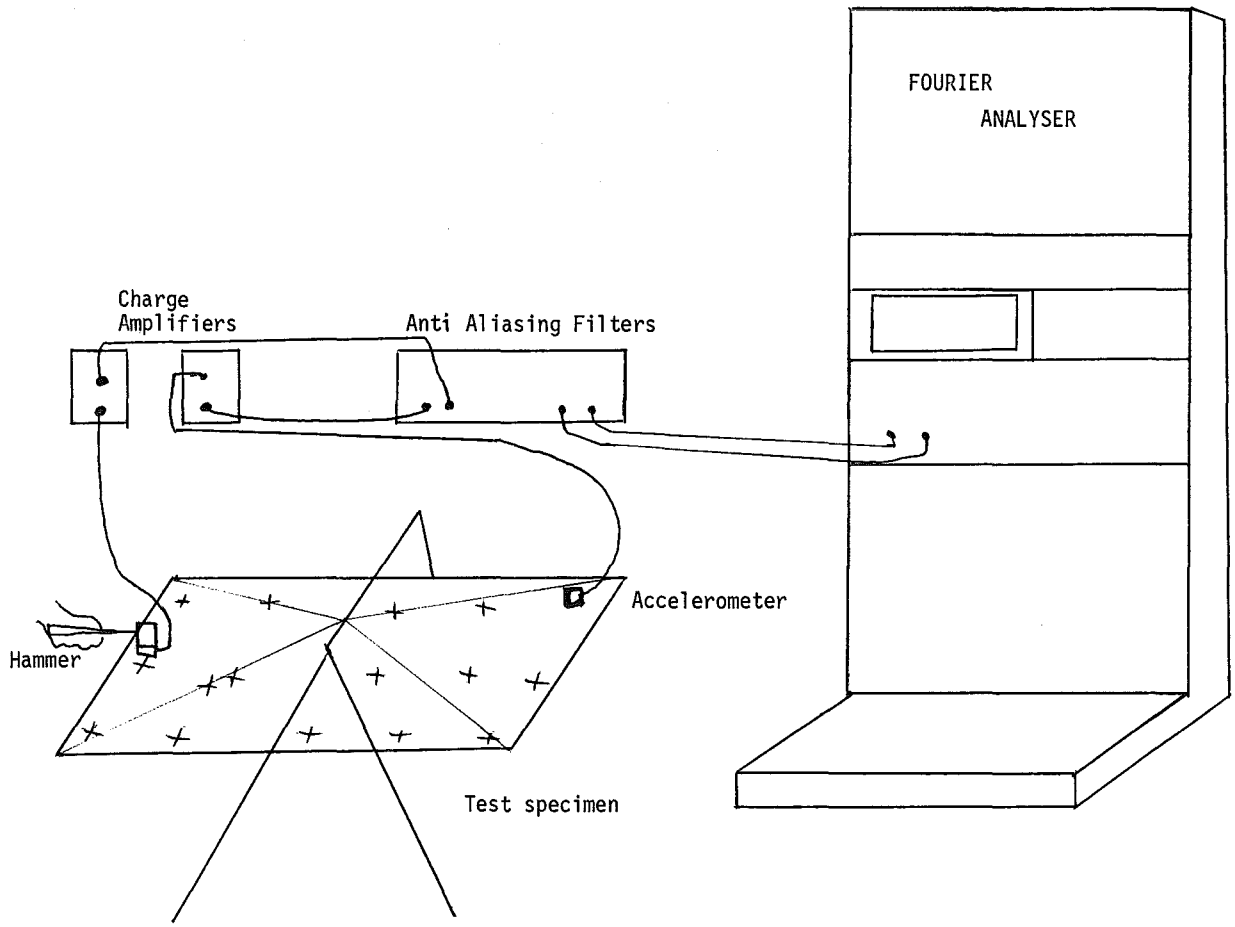


Fig. 3 HAMMER TEST SETUP

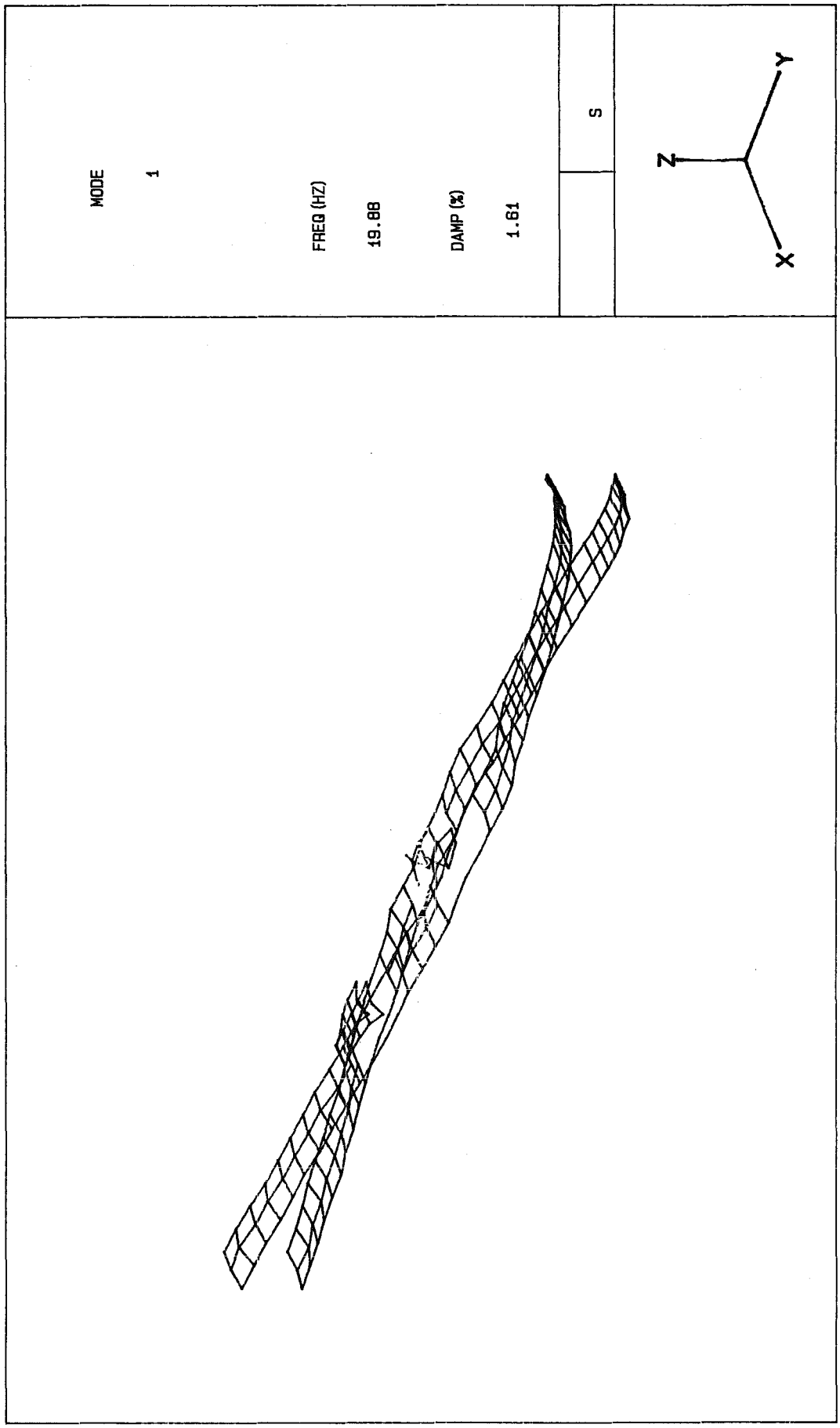


Fig. 4 THE FIRST BENDING MODE OF A SKI

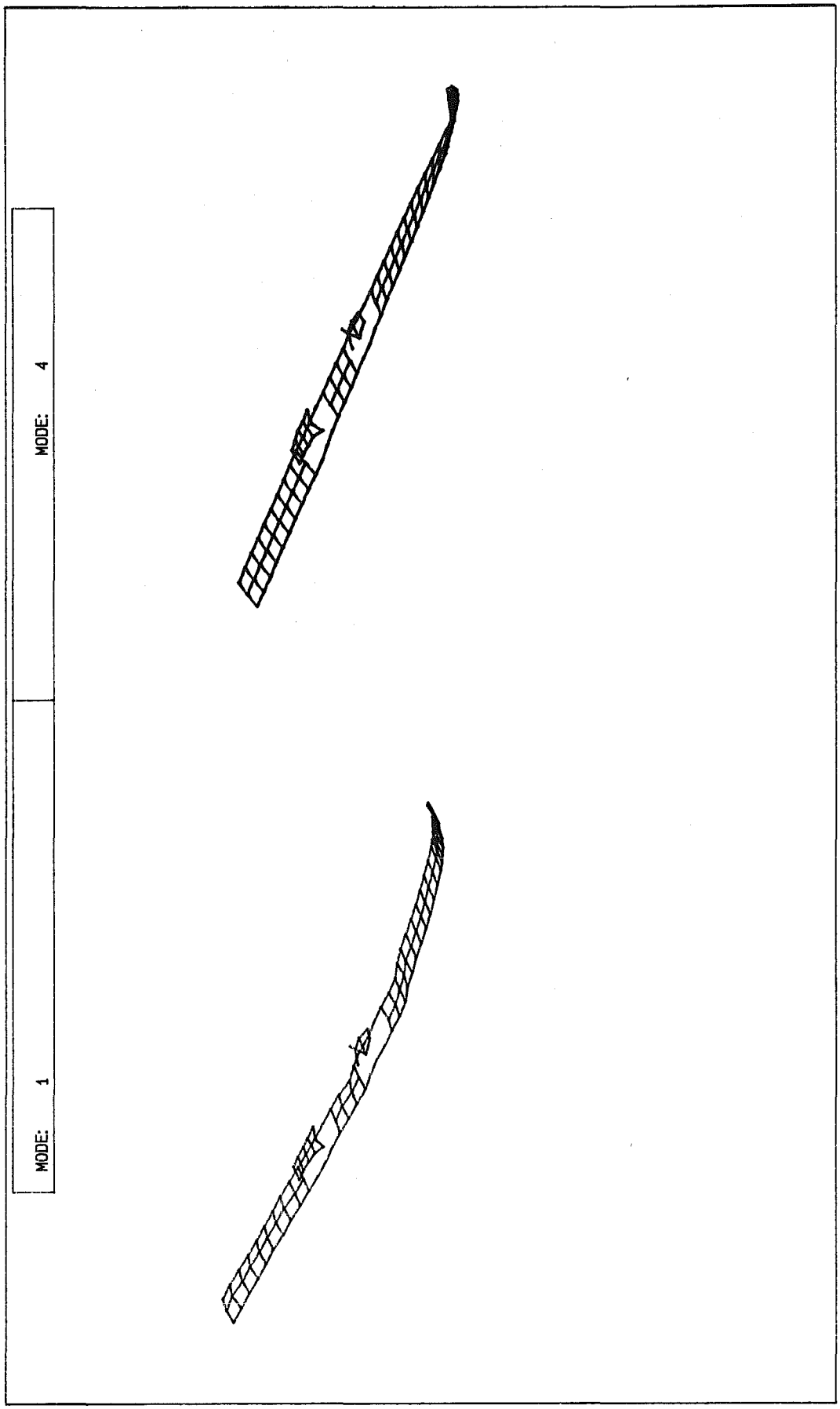


Fig. 4bis BENDING AND TORSIONAL MODES OF A SKI

REPORT DOCUMENTATION PAGE			
1. Recipient's Reference	**2. Originator's Reference**	**3. Further Reference**	**4. Security Classification of Document**
	AGARD-R-700	ISBN 92-835-0303-1	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 Rue Ancelle, 92200 Neuilly sur Seine, France

6. Title

MODERN DATA ANALYSIS TECHNIQUES IN NOISE AND VIBRATION PROBLEMS

7. Presented at

the von Kármán Institute, Rhode-St-Genèse, Belgium on 7–11 December 1981.

8. Author(s)/Editor(s)	**9. Date**
Various	November 1981
10. Author's/Editor's Address	**11. Pages**
Various	164 Pages

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Cover of all AGARD publications.

13. Keywords/Descriptors

Acoustics	Vibration
Noise (sound)	Data processing

15. Abstract

Aeroacoustics and hydroacoustics have many points in common when it comes to consideration of the characteristic features of far field radiation from acoustic or vibrating sources. The approach used to characterize such sources may be different, but in many cases this difference is more apparent than real, though it may be accentuated by the use of special-purpose instrumentation.

Data analysis techniques used by acoustics specialists on the one hand and vibration specialists on the other are analogous in many respects, even though data interpretation is made in response to different needs. In this special course specialists in the fields of acoustics, vibrations (in air and in water) and data analysis present their points of view. Particular emphasis is placed on points of similarity and on probable future developments.

The course was sponsored by the Fluid Dynamics and the Structures and Materials Panels of AGARD.

AGARD Report No. 700 Advisory Group for Aerospace Research and Development, NATO MODERN DATA ANALYSIS TECHNIQUES IN NOISE AND VIBRATION PROBLEMS Published November 1981 164 Pages Aeroacoustics and hydroacoustics have many points in common when it comes to consideration of the characteristic features of far field radiation from acoustic or vibrating sources. The approach used to characterize such sources may be different, but in many cases this difference is more apparent than real, though it may be accentuated by the use of special purpose instrumentation. P.T.O.	AGARD-R-700 Acoustics Noise (sound) Vibration Data processing	AGARD Report No. 700 Advisory Group for Aerospace Research and Development, NATO MODERN DATA ANALYSIS TECHNIQUES IN NOISE AND VIBRATION PROBLEMS Published November 1981 164 Pages Aeroacoustics and hydroacoustics have many points in common when it comes to consideration of the characteristic features of far field radiation from acoustic or vibrating sources. The approach used to characterize such sources may be different, but in many cases this difference is more apparent than real, though it may be accentuated by the use of special purpose instrumentation. P.T.O.	AGARD-R-700 Acoustics Noise (sound) Vibration Data processing
AGARD Report No. 700 Advisory Group for Aerospace Research and Development, NATO MODERN DATA ANALYSIS TECHNIQUES IN NOISE AND VIBRATION PROBLEMS Published November 1981 164 Pages Aeroacoustics and hydroacoustics have many points in common when it comes to consideration of the characteristic features of far field radiation from acoustic or vibrating sources. The approach used to characterize such sources may be different, but in many cases this difference is more apparent than real, though it may be accentuated by the use of special purpose instrumentation. P.T.O.	AGARD-R-700 Acoustics Noise (sound) Vibration Data processing	AGARD Report No. 700 Advisory Group for Aerospace Research and Development, NATO MODERN DATA ANALYSIS TECHNIQUES IN NOISE AND VIBRATION PROBLEMS Published November 1981 164 Pages Aeroacoustics and hydroacoustics have many points in common when it comes to consideration of the characteristic features of far field radiation from acoustic or vibrating sources. The approach used to characterize such sources may be different, but in many cases this difference is more apparent than real, though it may be accentuated by the use of special purpose instrumentation. P.T.O.	AGARD-R-700 Acoustics Noise (sound) Vibration Data processing

Data analysis techniques used by acoustics specialists on the one hand and vibration specialists on the other are analogous in many respects, even though data interpretation is made in response to different needs. In this special course specialists in the fields of acoustics, vibrations (in air and in water) and data analysis present their points of view. Particular emphasis is placed on points of similarity and on probable future developments.

The course was sponsored by the Fluid Dynamics and the Structures and Materials Panels of AGARD for the Special Course presented at the von Kármán Institute, Rhode-St-Genèse, Belgium on 7–11 December 1981.

ISBN 92-835-0303-1

Data analysis techniques used by acoustics specialists on the one hand and vibration specialists on the other are analogous in many respects, even though data interpretation is made in response to different needs. In this special course specialists in the fields of acoustics, vibrations (in air and in water) and data analysis present their points of view. Particular emphasis is placed on points of similarity and on probable future developments.

The course was sponsored by the Fluid Dynamics and the Structures and Materials Panels of AGARD for the Special Course presented at the von Kármán Institute, Rhode-St-Genèse, Belgium on 7–11 December 1981.

ISBN 92-835-0303-1

Data analysis techniques used by acoustics specialists on the one hand and vibration specialists on the other are analogous in many respects, even though data interpretation is made in response to different needs. In this special course specialists in the fields of acoustics, vibrations (in air and in water) and data analysis present their points of view. Particular emphasis is placed on points of similarity and on probable future developments.

The course was sponsored by the Fluid Dynamics and the Structures and Materials Panels of AGARD for the Special Course presented at the von Kármán Institute, Rhode-St-Genèse, Belgium on 7–11 December 1981.

ISBN 92-835-0303-1

Data analysis techniques used by acoustics specialists on the one hand and vibration specialists on the other are analogous in many respects, even though data interpretation is made in response to different needs. In this special course specialists in the fields of acoustics, vibrations (in air and in water) and data analysis present their points of view. Particular emphasis is placed on points of similarity and on probable future developments.

The course was sponsored by the Fluid Dynamics and the Structures and Materials Panels of AGARD for the Special Course presented at the von Kármán Institute, Rhode-St-Genèse, Belgium on 7–11 December 1981.

ISBN 92-835-0303-1

AGARD
NATO OTAN
7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE
FRANCE
Telephone 745.08.10 · Telex 610176

DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres, but if not may be purchased in Microfiche or Photocopy form from the Purchase Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD – VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Defence Science Information Services
Department of National Defence
Ottawa, Ontario K1A OK2

DENMARK
Danish Defence Research Board
Østerbrogades Kaserne
Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92320 Châtillon sous Bagneux

GERMANY
Fachinformationszentrum Energie,
Physik, Mathematik GmbH
Kernforschungszentrum
D-7514 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force General Staff
Research and Development Directorate
Holargos, Athens

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3, Piazzale Adenauer
Roma/EUR

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 126
2600 A.C. Delft

NORWAY
Norwegian Defence Research Establishment
Main Library
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Direcção do Serviço de Material
da Forca Aerea
Rua da Escola Politécnica 42
Lisboa
Attn: AGARD National Delegate

TURKEY
Department of Research and Development (ARGE)
Ministry of National Defence, Ankara

UNITED KINGDOM
Defence Research Information Centre
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RE

UNITED STATES
National Aeronautics and Space Administration (NASA)
Langley Field, Virginia 23365
Attn: Report Distribution and Storage Unit

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

PURCHASE AGENCIES

Microfiche or Photocopy
National Technical
Information Service (NTIS)
5285 Port Royal Road
Springfield
Virginia 22161, USA

Microfiche
Space Documentation Service
European Space Agency
10, rue Mário Nikis
75015 Paris, France

Microfiche
Technology Reports
Centre (DTI)
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RF
England

Requests for microfiche or photocopies of AGARD documents should include the AGARD serial number, title, author or editor, and publication date. Requests to NTIS should include the NASA accession report number. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical
Information Facility
Post Office Box 8757
Baltimore/Washington International Airport
Maryland 21240, USA

Government Reports Announcements (GRA)
published by the National Technical
Information Services, Springfield
Virginia 22161, USA

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London W1P 1HD

ISBN 92-835-0303-1